Towards the establishment and implementation of an institutional repository at the University of Cape Town (UCT): a case study

A major dissertation submitted in full fulfilment of the requirements for the award of the degree of Master of Philosophy (M.Phil) in Library and Information Science

by

Annah Sephene Macha (MCHANN004)

Supervisor: A/Professor Karin de Jager

Department of Information and Library Studies

Faculty of Humanities

University of Cape Town

April: 2012

DECLARATION

This work has not been previously submitted in whole or in part, for award of any degree. It is my own work. Each contribution to, and in quotation in, this dissertation from the work or works of other people has been attributed and has been cited and referenced.

PLAGIARISM DECLARATION:

1. I know that plagiarism is wrong. Plagiarism is to use another's work and to pretend that it is one's own.
2. I have used the augmented Harvard convention for citation and referencing. Each significant contribution to, and quotation in, this easy/report/project..... from the work, or works, of other people has been attributed, and has been cited and referenced.
3. This essay/report/project.....is my own work.
4. I have not allowed, and will not allow, anyone to copy my work with the intention of passing it off as his or her own work.

SIGNATURE: _____

DATE OF SUBMISSION:_____2012

ACKNOWLEDGEMENTS

Firstly, I would like to acknowledge the blessings of the Lord Almighty for His protection, wisdom and good health bestowed upon me during the period of my studies. I also wish to express my sincere gratitude to the Department of Teacher Training and Development (TT&D), under the Teaching Service Management (TSM) of the Ministry of Education (Botswana), for the opportunity, they granted me through their sponsorship to pursue my study at the University of Cape Town (UCT).

I greatly acknowledge the support and encouragement given to me by my supervisor, Associate Professor Karin de Jager, and deeply appreciate her dedication, patience, guidance and love she provided me throughout the study. She taught me to be confident about my work and my views. Without her critical and constructive advice, I would not have reached this far. Therefore for me *to have seen this far is by standing on her shoulders* (Isaac Newton).

I gratefully recognise the outstanding support given to me by Janine Dunlop and her co-workers from the Digital Initiative Unit in the UCT Libraries. She exposed me to digitisation, its terminologies and the practical aspects of it. I would be ungrateful if I do not acknowledge the Manager of the Manuscripts and Archiving Department, Lesley Hart, and her co-workers for their cooperation and support during my interviews. Without them, I would not have completed this research. I also deeply appreciate the work of Linda Anne Fraenkel Kelly who sacrificed her time to edit my work.

I also express my profound thanks and appreciation to people who rendered their spiritual and moral support that made this dream a reality. These are my pastor, counsellor, guardian, and father figure, Reverend Bishop Jeremiah Toteng and his wonderful family for their everlasting love, encouragement and guidance; my dearest friends, Mr. Kgaogano Sefho, and his loving wife, Fikile Sefho; and my aunts, Mrs Catherine Mbeha and Mrs Esther Oaitse Tshukudu, for their support, love and prayers.

Last but not the least, I wish to appreciate my wonderful brother and friend, Sinka Matengu who was always there for me and gave me so much moral support, and inspirations during the course of my studies.

ABSTRACT

The concepts of open access and scholarly publishing are still gaining momentum in Africa, especially South Africa. Increasingly, institutional repositories are being planned and developed by universities throughout the world especially in the first world countries, which have taken the lead. Institutional repositories have developed because of changes in scholarly communication where journal prices are high and libraries are finding it difficult to subscribe to them. Communication technology in the form of the internet brought a solution to the problem. Researchers, authors and libraries now advocate for the open access model of scholarly communication. This study explores the developments associated with the establishment and implementation of an Institutional Repository at UCT. By establishing an institutional repository, UCT aims to be visible by digitizing its resources as a contribution to African Scholarship and as an incentive for digital collaboration on the continent, and as a showcase for its research. This is compared to the institutional repository at the University of Pretoria (UP) in order to determine best practices from both institutions. Factors for a successful institutional repository will be discussed and used to design benchmarking assessments against which the UCT Institutional Repository will be evaluated.

TABLE OF CONTENTS

Preliminary Pages ..... i- xvi

Title page .....

Declaration i

Acknowledgements ii

Abstract iii

List of abbreviations ..... xiii-xvi

Chapter 1: Introduction 1

1.1 Introduction 1

1.2 Conceptual Framework 2

1.3 Background of the study 3

1.3.1 Towards developing an institutional repository at the University of Cape Town 4

1.4 Aims and objectives of the study 5

1.4.1 Research objectives 5

1.5 Primary research questions 6

1.5.1 Secondary research questions 7

1.6 A comparative analysis of UCT and UP repositories 7

1.7 Significance of the study 8

1.8 Limitations of the study 8

1.9 Outline of chapters 9

1.10 Conclusion 9

Chapter 2: Institutional Repositories: Theory and Practice 11

2.1 Introduction 11

2.2Historical background of institutional repositories 11
2.2.1 The concept of institutional repositories 12
2.2.2 Reasons for emergence of institutional repositories 13
2.2.3 The nature of content that institutional repositories collect 14
2.2.4 Responsibility for depositing materials in the repository 15
2.3 Exploration, review and analysis of related concepts 16
2.3.1 Digital libraries 16
2.3.2 Digital information 17
2.3.3 Digitisation 17
2.3.4 Curation 18
2.3.5 Archiving 18
2.3.6 Digital electronic archiving.....19
2.3.7 Preservation.....19
2.3.8 Interoperability 19
2.3.9 Open access institutional repositories 20
2.3.10 Data and metadata 20
2.4 Digital Preservation.....21
2.4.1 Digital preservation challenge 22
2.4.2 Technology obsolescence 23
2.4.3 Preservation policy.....24
2.4.4 The OAIS model of preservation.....25
2.5 Academic libraries as centres for research support 25
2.5.1 The role played by libraries in preservation, curation and archiving 27
2.6 Scholarly communications 28
2.6.1 Stages of scholarly communications 29

2.6.2 Electronic publishing and publication 30
2.7 Theoretical Framework for institutional repositories 31
2.7.1 Exploring the Open Access model of publication 31
2.7.2 Policy and principle statements about Open Access 32
2.7.3 Bermuda Principles 1966 32
2.7.4 Budapest Open Archive Initiative (BOAI) - 2002 32
2.7.5 Bethesda Statement on Open Access publishing - 2003 33
2.7.5.1 Bethesda Statement 33
2.7.5.2 Berlin Declaration - 2003 34
2.8 Degrees of Open Access 34
2.8.1 Routes to the publications on Open Access journals 35
2.8.1.1 Green route 35
2.8.1.2 Gold route 35
2.8.1.3 Types of Open Access publication 36
2.8.2 Disciplinary digital repositories 36
2.8.3 Departmental repositories or personal homepages 37
2.8.4 Electronic Theses and Dissertations digital repositories 37
2.9 Factors for establishing a successful institutional repository 38
2.9.1 Identify important role players 38
2.9.2 Address issues of resources 39
2.9.3 Evaluate software 39
2.9.4 Formulate policy for the institutional repository 39
2.9.5 Restructuring the library to accommodate change 41
2.9.6 Licensing 41
2.10 Criteria for a successful institutional repository 42

2.10.1 Content 44
2.10.2 Use 44
2.10.3 Submissions 45
2.10.4 Advocacy 46
2.10.4.1 Addressing authors concerns 47
2.10.4.2 Pillars of advocacy 49
2.10.5 Support 51
2.10.6 Influence 52
2.10.7 Interoperability 52
2.11 Proposals for dealing with securing contents for an institutional repository 53
2.11.1 Self-archiving 53
2.11.2 Encourage deposit of staff publication 54
2.11.3 Policies that mandate deposit 54
2.11.4 Marketing the repository 55
2.11.4.1 Knowing the audience 56
2.11.4.2 Content of the marketing message 56
2.11.4.3 Delivering the message 56
2.11.5 Barriers of content recruitment for an institutional repository 57
2.12 Software used for institutional repositories 57
2.13 International developments on institutional repositories 60
2.13.1 Institutional repositories in the UK 60
2.13.1.1 Institutional repository challenges faced in the UK 63
2.13.2 Institutional repositories in the USA 63
2.13.2.1 Institutional repository challenges faced in the USA 65
2.13.3 Institutional repositories in Canada 65

2.13.3.1 Institutional repository challenges faced in Canada 66
2.13.4 Institutional repositories in Australia.....67
2.13.4.1 Institutional repository challenges faced in Australia 68
2.13.5 Institutional repositories in Africa 69
2.13.5.1 Institutional repository challenges faced in Africa 70
2.13.5.2 The status of institutional repositories in South Africa 71
2.13.5.3 Repositories in South Africa 2011 72
2.14 Metadata assignment in a digital institutional repository 73
2.14.1 The importance of metadata in repository 74
2.14.2 Metadata creation 75
2.14.3 Metadata use 75
2.14.4 Metadata standards 76
2.14.5 Types of metadata 77
2.14.6 Important points to note about metadata.....78
2.15 Institutional repository and the copyright issues 78
2. 16 Challenges in establishing institutional repositories in Africa 80
2.16.1 Inadequate funding.....80
2.16.2 Inadequate power and ICT connectivity and infrastructure.....80
2.16.3 Copyright and intellectual property right issues 82
2.16.4 Lack of advocacy 82
2.16.5 Lack of faculty involvement leads to low content.....83
2.16.6 Technological barriers 83
2.16.7 Institutional culture, and politics 83
2.17 Institutional Repositories and the future.....84
2.18 Conclusion.....86

Chapter 3: Research Methodology and Design **88**

3.1 Introduction 88

3.2 Primary research questions 88

3.2.1 Secondary research questions 88

3.3 Qualitative Research methodology 89

3.3.1 Case study 91

3.3.1.1 Argument against the use of case studies 92

3.3.1.2 Types of case studies 93

3.3.1.3 Reliability and validity of the study 93

3.4 Research design 94

3.5 Population of the study 96

3.6 Data collection techniques 97

3.6.1 Documents 97

3.6.2 Archival records 98

3.6.3 Direct observation 99

3.6.4 Participation observation 99

3.6.5 Interviews 100

3.6.6 Physical artefacts 102

3.6.7 Three guideline of case study data collection 102

3.7 Data analysis 103

3.8 Conclusion 104

Chapter 4: Establishing an I.R at UCT: a case study **106**

4.1 Introduction 106

4.2 Towards developing an Institutional Repository at UCT 106

4.3 How the IR was established in the UCT libraries 107

4.3.1 Contents of the UCT Institutional Repository 108
4.3.2 Digital collections 109
4.3.3 Finding aids 111
4.3.4 Electronic theses and dissertations 113
4.4 The case study at UCT 113
4.5 Case data 114
4.5.1 Participation observation 114
4.5.2 Physical artefacts 116
4.5.3 Administrative documents 117
4.5.4 Interviews 119
4.6 Themes from the interviews 120
4.6.1 Responsibility 120
4.6.2 Motivation for establishing an institutional repository 120
4.6.3 Stakeholders and resources 121
4.6.4 Planning 122
4.6.5 Open access initiative 122
4.6.6 Consolidation of departmental repositories 123
4.6.7 Institutional repositories assessment 123
4.6.8 Other issues that emerged 124
4.6.9 Challenges that were encountered 125
4.6.9.1 Shortage of staff 125
4.6.9.2 Shortage of hardware 126
4.6.9.3 Lack of flexibility of DigiTool 127
4.6.9.4 The problem of archiving audio visual materials 127
4.6.9.5 Insufficient training 128

4.6.9.6 Office Space 129
4.6.9.7 Other issues 129
4.6.7 Resolution of challenges 129
4.8 Conclusion 131
Chapter 5: Criteria for a successful IR: a comparison between UCT and UP 132
5.1 Introduction 132
5.1.1 Background of the University of Pretoria 132
5.1.2 Establishment of an institutional repositories at UP 133
5.2 A comparative analysis of the UCT and UP institutional repositories 135
5.2.1 Identify important role players 136
5.2.2 Address issues of resources 136
5.2.3 Evaluate software 137
5.2.4 Formulate Policy for institutional repositories 139
5.2.Restructure the library to accommodate change 140
5.2.6 Licensing 141
5.3 How UCT IR complies with the criteria for a successful institutional repository 142
5.3.1 Content 142
5.3.2 Use 142
5.3.3 Submissions 143
5.3.4 Advocacy 144
5.3.5 Support 144
5.3.6 Influence 145
5.3.7 Interoperability 146
5.4 A critical summary of UP as a benchmark and way forward for UCT 146
5.5 Conclusion 148

Chapter 6: Discussion of key findings, conclusion and recommendations 149

6.1 Introduction 149

6.2 Discussion of key findings 149

6.2.1 Content 149

6.2.2 Use 150

6.2.3 Submission 150

6.2.4 Advocacy 151

6.2.5 Support 151

6.2.6 Influence 151

6.2.7 Interoperability 152

6.3 Summary of research questions and findings that answer the questions 152

6.4 Recommendations 154

6.4.1 Ensuring that the unit moves to a bigger location 154

6.4.2. Employ more members of staff 154

6.4.3. Buy more equipment 154

6.4.4 In-service training 155

6.4.5 Market the repository 155

6.5 Overall conclusion 155

6.6 Suggestions for further research 156

References 158

Appendices 177

Appendix A:Permission letter for conducting research 177

Appendix B: Administrative documents 179

Appendix C: Interview guide questions 180

Appendix D:UCT Libraries- Scale of charges for reprographic and other services 2011 ...185

LIST OF ABBREVIATIONS

AAU	Association of American Universities
ACRL	Association of College and Research Libraries
ADT	Australian Digital Theses Program Expansion and Redevelopment
AHDS	Arts and Humanities Data Service
APSR	Australian Partnership for Sustainable Repositories
ARL	Association of Research Libraries
ARROW	Australian Research Repositories Online to the World
ARWU	Academic Ranking of World Universities
ASSAf	Academy of Science of South Africa's Journal
BD	Berlin Declaration
BOAI	Budapest Open Access Initiative
CARL	Canadian Association of Research Libraries
CCA	Center of Curating the Archive
CIHR	Canadian Institutes of Health Research
CNI	Coalition for Networked Information
CSIR	Council of Scientific and Industrial Research
CURL	Consortium of University Research Libraries
DARPR	Defense Advanced Research Projects Agency
DATAD	Database of African Theses and Dissertations

DCMI	Dublin Core Metadata Initiative
DIU	Digital Initiatives Unit (in the UCT libraries)
DEST	Department of Education, Science and Training (in Australia)
DINI	Deutsche Initiative for Networked Information
DISA	Digital Imaging South Africa
DRIVER	Digital Repository Infrastructure Vision for European Research
DST	Department of Science and Technology
eIFL	Electronic Information for Libraries
eSCFs	electronic Scholarly Communication Forums
ETD	Electronic thesis and dissertations
FAIR	Focus on Access to Institutional Resources
FEDORA	Flexible Extensible Digital Object and Repository Architecture
FRODO	Federated Repositories of Digital Objects
IPR	Intellectual Property Rights
IR	Institutional Repository
ISBN	International Standard Book Number
JISC	Joint Information Systems Committee
KNUST	Kwame Nkrumah University of Technology (In Ghana)
MAMS	Meta Access Management System
MIRACLE	Making Institutional Repositories a Collaborative Learning Environment
MIT	Massachusetts Institute of Technology

Mss & A	Manuscripts and Archives Department- at University of Cape Town
NDLTD	Networked Library of Theses and Dissertations
NSF	National Science Foundation
OA	Open Access
OAI	Open Archive Initiative
OAI - PMH	Open Archive Initiative the Protocol for Metadata Harvesting
OAIS	Open Archival Information System
OER	Open Educational Resources
OpenDOAR	Directory of Open Access Repositories
QS	Quacquarelli Symonds
RCUK	Research Council of United Kingdom
ROAR	Registry of Open Access Repositories
RoMEO	Rights Metadata for Open Archiving
SA	South Africa
SARI	South African Research Information Services
SARUA	Southern African Regional Universities Association
SHERPA	Securing a Hybrid Environment Research Partnerships and Access
SPARC	Scholarly Publishing and Academic Resources Coalition
SWORD	Simple Web Service Offering Repository Deposit
UCT	University of Cape Town
UK	United Kingdom

UNESCO	United Nations Educational, Scientific and Cultural Organization
UP	University of Pretoria
UPeTD	University of Pretoria electronic theses and dissertation repository
USA	United States of America
WWW	World Wide Web

CHAPTER ONE

1.1 Introduction

Developments in Information and Communication Technology (ICT) and the advent of the internet have significantly influenced the way libraries disseminate information to users. The internet has resulted in the emergence of open 2 institutional repositories, which are digital archives for the preservation and dissemination of institutional research outputs found in academic libraries (Christian, 2008). The growth of open access institutional repositories is still in its infancy in many developing countries but is well established in developed countries. Academic and research institutions in many developing countries in sub-Saharan Africa are still struggling to overcome many challenging issues in an attempt to make their research outputs openly accessible by means of establishing institutional repositories (Christian, 2008).

Establishing and maintaining repositories has created an opportunity for African countries to participate in eResearch (Van Deventer and Pienaar, 2008). eResearch refers to "the use of advanced information and communication technologies to do research and do research collaboration better" (Monash University eResearch Centre website, 2011). It enables researchers to draw on perspectives and resources from a range of participants by using technology to develop new insights and new solutions to complex problems. Thus, eResearch uses technology as a facilitator to researcher collaboration. The Monash eResearch Center (2011) explains that eResearch enhances existing research processes making them more efficient and effective. Researchers around the developed world have been exploiting the opportunities of eResearch in the last ten years, mainly in countries such as the United Kingdom (UK) and the United States of America (USA). African countries establishing and maintaining institutional repositories are creating opportunities to participate in eResearch, exploiting the opportunity to make their knowledge output more widely known and accessible through the internet.

Van Der Merwe (2008) indicated that the challenge facing most organizations is managing the vast volume of knowledge artefacts. This ranges from research reports, books, research papers and conference papers to *ad hoc* explicit records in a variety of electronic formats and audiovisual material. She proposes that the effective management, preservation and curation of

the knowledge artefacts are required to facilitate efficient retrieval. The emergence of institutional repositories is the result of an attempt to address some of these demands.

Following developments in South Africa regarding the establishment of institutional repositories, this study investigates developments towards the establishment and implementation of the institutional repository at UCT as a possible participant in eResearch. The study will show how the institutional repository at UCT can address the demands of digital preservation, curation and the ability to create and transfer knowledge.

1.2 Conceptual Framework

Broadly, the main concerns of this study are to explore the challenges of establishing and implementing a digital institutional repository. According to Kim (2007), institutional repositories can be conceptualized as comprising the following three entities: electronic scholarly communication forums, digital libraries and knowledge management systems. Several studies explain that an institutional repository would be a new strategy for facilitating changes in electronic scholarly communication (Chan, 2004; Crow, 2002; Lynch, 2003; Shearer, 2003). Therefore, the institutional repository can be embraced into a broad family of electronic scholarly communication forums. Below is an explanation of the conceptualizations mentioned above.

The first conceptualization of institutional repositories as electronic scholarly communication forums means that institutional repositories follow the socio-technical network model. This model, as explained by Kim (2007), combines participants with different roles, responsibilities and resource flows with new networking technologies. Institutional repositories in this model are seen as platforms that bring together the social and technical elements. The social element refers to the interaction of the people who manage institutional repositories and those who contribute contents in the repository. The technical aspects refer to those dealing with the technology used to make the contents available and accessible to users.

The second conceptualization of institutional repositories as digital libraries also follows the socio-technical network model mentioned above, but also brings together networks of

technology, people and practices and how these three interact (Kim, 2007). Institutional repositories are viewed not only as information retrieval systems or as collections of digital resources from their communities, but also as information institutions like libraries, archives and museums (Kim, 2007). An institutional repository can be classified as a type of digital library because, while it houses a collection of digital images, documents, audiovisual materials and datasets, it also ensures that these collected items are archived, managed, and preserved to facilitate long-term availability and easy retrieval. An institutional repository can be established by any organisation, for example, Science Councils, Non-Governmental organizations and universities through contributions from its community of scholars, students and stakeholders.

The third perspective views institutional repositories as knowledge management systems. This means that institutional repositories are information technology based systems. They are developed and established to support and enhance the institution's processes of knowledge creation, storage, retrieval, transfer, dissemination and use. Kim (2007) notes that an increasing number of faculty members and students at a university utilize information technology not only to access information, but also to create new intellectual research output in digital forms. The approach to knowledge management, as viewed by Kim (2007), is relevant to the implementation of institutional repositories because they manage a wide range of digital information created at a university.

In this dissertation the concept of institutional repository is understood to refer to electronic scholarly communication systems and digital libraries, as will be discussed below.

1.3 Background to the study

South Africa's oldest university is the University of Cape Town (UCT). It was founded in 1829 "as a school for boys that also provided some tertiary education. It was then known as the South African College. It did not admit women, they were only admitted to the college after 1887. UCT was formally established as a university in 1918" (SARUA website). UCT is one of Africa's leading teaching and research institutions. According to the 2011 Academic Ranking of World Universities (ARWU) it is ranked number 255, and number one in Africa. The ARWU rankings, show that UCT has been in the first position in the continent for over five years. Another ranking organisation, the Quacquarelli Symonds (QS) World University Rankings for 2011/12, ranks

UCT at position 156. The QS system states that to date UCT is the only African university in the top 200s. The UCT Website, states that UCT has "more than 60 specialist research units that provide supervision for postgraduate work and is home to more than a quarter of South Africa's A-rated researchers or academics who are considered world leaders in their fields (UCT 1)." If the institution produces so much research output, it is reasonable to assume that it requires an institutional repository that will preserve, manage, share and house this intellectual property.

Generally, African universities are ranked lowest in terms of research output as compared to other universities in the world (Moahi, 2009). She further states that this is something that institutional repositories can address. She further notes that though UCT, is ranked as number one in Africa and yet ranked at 359 out of 6000 universities of the world. This indicates that the online presence of African universities is very low. The UCT Institutional Repository will add to the number of repositories in Africa and increase the online presence of African research.

1.3.1 The development of an institutional repository at the University of Cape Town

According to Dunlop and Hart (2005), the University of Cape Town Libraries identified a collection of photographs of the San people which had been listed and registered by UNESCO as documentary in its Memory of the World register and started to digitise it. This was a project undertaken by the Manuscripts and Archives Department in 2001 on a small scale. The project mainly focused on digitizing photographs of the San people, which were taken between the years 1910 and 1920s by Dorothea Bleek. According to Dunlop and Hart (2005) also quoted by Macha and De Jager (2011), the San photographic collection is an important photographic collection at UCT as it attracts researchers around the world.

The Manuscripts and Archives Department (Mss & A) of the University of Cape Town project, was made possible by the attendance of the UCT's team to workshops conducted by Digital Imaging South Africa (DISA). These workshops taught the participants international guidelines, overviews of the processes, and capture parameters used when digitizing images. It also helped them with lessons on the types of metadata used and the overview of the processes of digitization which were the by the department as they carried out the project.

Other digitization projects that followed at UCT were instigated by individual departments. In 2003 the Department of Computer Science started to digitise theses and dissertations of their

master's and doctoral students. This was followed by the Faculty of Law in 2005, which established the LawSpace Digital Subject repository. Below is a list of repositories found in UCT as adopted from Macha and De Jager (2011: 2)

1. "Faculty of Law - UCT LawSpace
2. Department of Computer Science - UCT Computer Science Research Document Archive
3. Manuscripts and Archives Department, and
4. Open Educational Resources (OER) in UCT. Open Content is not a repository but a directory for educational materials. It is the web portal for accessing open teaching and learning content from UCT. It is produced by OER project in the Centre for Educational Technology at UCT. The directory aims to display the teaching materials of UCT academics and to encourage the publication of open resources."

The above repositories will be brought together as UCT establishes its institutional repositories so that they can be just one repository for the institution.

1.4 Aims and objectives of the study

In view of the above discussion, this study aims to:

- Explore the steps followed in the implementation, development and establishment of an institutional repository and the measures taken to overcome challenges.
- Compare developments of the institutional repository at UCT with that at University of Pretoria (UP).
- Explore the extent to which the process followed at UCT has lead to a resource that meets with the criteria for success.

1.4.1 Research Objectives

In order to meet the above aims, the following objectives were identified:

- To explore factors that lead to the establishment of an institutional repository at UCT libraries.

- To establish what resources and structures are involved when developing a institutional repository.
- To investigate to what extent the institutional repository is an open access initiative.
- To establish if the Digital Initiatives Unit is involved in bringing the departmental repositories together as the UCT Institutional Repository.
- To establish the criteria that could be used to evaluate the success of an institutional repository.
- To identify challenges encountered by the Digital Initiatives Unit in the establishment of their institutional repository.
- To explore how challenges may be overcome.
- To compare the developments of the UCT Institutional Repository with that at UP.
- To evaluate the developmental success of the UCT institutional repository.

1.5 Primary Research Questions

From the above objectives and aims, the researcher derived the following primary and secondary research questions. The methodology chosen for this study on the development, establishment and implementation of institutional repositories takes into consideration the research questions:

- What are the steps involved in the development, establishment and implementation of an institutional repository and what measures have to be taken to overcome identified challenges?
- How does the UCT institutional repository compare and contrast with that at UP?
- To what extent will the process followed at UCT in developing an institutional repository lead to a resource that meets with criteria for success?

1.5.1 Secondary Research Questions

- What are the factors that influenced UCT to establish an institutional repository?
- Which resources and structures have been put in place in developing the institutional repository at UCT?
- To what extent is the repository going to be an open access initiative?
- Are the existing departmental digital repositories going to be brought together under the University of Cape Town Institutional Repository?
- What are the criteria that could be used to evaluate the success of an institutional repository?
- What challenges have been encountered in the establishment of the UCT institutional repository?
- How were these challenges overcome?
- To what extent does the newly established UCT Institutional Repository meet the criteria of success?

1.6 A comparative analysis of UCT and UP repositories

In addition to documenting the developments and events that lead to the implementation and establishment of the UCT Institutional Repository and the challenges it faced, this study will include a comparative analysis of the institutional repositories at UCT and UP. The study will investigate the factors that UP considered when implementing and establishing its institutional repository as well the challenges faced when executing such an endeavor. How these challenges were overcome would be beneficial to other institutions who wish to develop a repository. Therefore factors will be extrapolated for benchmarking from both university's experiences as well as available literature in order to compare the institutional repositories at UCT and UP. The motivation for this comparison is that the University of Pretoria is one of the first universities in South Africa to establish an institutional repository and it helped the CSIR to establish its own. This is a sign that it is more informed about setting up a repository and could be used as a bench-mark by other institutions wishing to establish such a project.

1.7 Significance of the study

Institutional repositories are a new development that attempts to make research outputs openly accessible through the internet. They could help academic staff to share their research internationally with others in the same field and with those who have similar interests.

The study will contextualise the development and establishment of the institutional repository at UCT and UP. Findings should enable the Digital Initiatives Unit at UCT to market the repository and gain a better understanding of how it is perceived, utilised and managed by the different departments involved.

1.8 Limitations of the study

The study will be confined to the selected departments at UCT listed below. They have been chosen because currently they have functional departmental digital repositories in UCT:

- Faculty of Law - UCT LawSpace,
- Department of Computer Science - UCT Computer Science Research Document Archive,
- Manuscripts and Archives Department and
- Open Educational Resources - UCT- Centre for Educational Technology.

The case study approach will be used. According to Yin (2009) the limitations of case study research are that scholars tend to view them as a less preferred form of inquiry compared to either experiments or survey research. This is because they are reputed not to follow systematic procedures and as such produce biased views of the findings and conclusions. Another limitation is that they lack flexibility because of the existence of many methodological texts, which provide the investigators with specific procedures to be followed. The most common limitation of case studies is that they provide little basis for scientific generalisation as their aim is to generalize theories (Yin, 2009). This is known as analytic generalisation (Yin, 2009). Case studies can only be generalisable to theoretical propositions and not to populations or universes. This means that case studies do not represent a "sample," and Tellis (1997) states that case studies are not sample

researches. The reasons for using the case study approach, as well as its positive aspects, are discussed in Chapter 3 (cf 3.3.1).

The other limitation is that the research on UP will be based on published work, as time will not permit the researcher to travel to Pretoria to conduct some interviews.

1.9 Outline of chapters

The following is a brief outline of the chapters:

- **Chapter 1**: Introduction, conceptual framework of the study, background of the research problem, statement of the research problem, aims and objectives of the research problem, research questions, significance of the study, and the limitations of the study, outline of chapters and conclusion.
- **Chapter 2**: Historical background of institutional repositories, exploration, review and analysis of related concepts, scholarly communication, open access and literature review and conclusion;
- **Chapter 3**: Research methodology, research design, population of the study and data collection instruments and conclusion;
- **Chapter 4**: Establishing an institutional repository at UCT: Case study, case data analysis, theme from interviews, challenges encountered and how they were overcome.
- **Chapter 5**: Criteria for a successful institutional repository: a comparative analysis of UCT and UP.
- **Chapter 6**: Discussion of the key findings based on criteria for a successful institutional repository, recommendations and conclusion.

1.10 Conclusion

This chapter briefly introduced the different departmental digital repositories at UCT. It outlined the research questions, sub questions and objectives of the study. The conceptual framework of the study was also discussed. The importance of the study, its limitations as well as the overview of the thesis' chapters is provided. As UCT establishes the institutional repository, its research

endeavours will become visible to the world, contributing towards world heritage and in the process achieving its goal of being Africa's leading research university.

The next chapter will explore, review and analyse related concepts, history and background literature on institutional repositories, scholarly communication and open access. It will also review African participation to eResearch by way of establishing institutional repositories, focusing on the South African situation, particularly at initiatives and challenges and how these can be overcome.

CHAPTER TWO

INSTITUTIONAL REPOSITORIES: THEORY AND PRACTICE

2.1 Introduction

As noted in chapter one, institutional repositories are becoming increasingly important to academic institutions and their academic communities. This chapter will firstly look at the historical background of institutional repositories. This will be followed by exploring the terms used within the study. The chapter will also explore the open access movement, which aims at making research accessible to the wider research communities through institutional repositories. New ways of scholarly communication brought about by the advancements in Information, and Communication Technologies (ICT) are discussed. The chapter concludes by reviewing the literature, focusing firstly on the first world countries and secondly on Africa, especially South Africa.

2.2 Historical background of institutional repositories

The Open Archives Initiative (OAI) meeting, held in Santa Fe, New Mexico in October 1999 boosted the movement for institutional repositories (Hitchcock et al., 2007; Van de Sompel & Lagoze, 2000). Later this was continued by the Open Source Software conference, and the meeting of the Budapest Open Access Initiative (BOAI) held during December 2001 in Hungary and the Berlin Declaration (BD) meeting held during October 2003 in Germany (Pappalardo & Fitzgerald, 2007). The BOAI (2007) emphasized the fact that the internet had a fundamental and dramatic influence on the availability and distribution of scholarly information. The BD later emphasized the BOAI statement, which stated that the internet should be seen as a tool for improving the "global scientific knowledge base". The BD also emphasized that all stakeholders should specify the measures that should be taken to make effective use of the internet (Bailey, 2006). This was the time when librarians began digitizing content in their special collections, making it accessible worldwide.

A growing body of literature regarding institutional repositories emerged around the year 2002 (Christian, 2008). The literature concentrated on major research universities in the United States

and United Kingdom which had implemented institutional repositories (Markey et al., 2007). A major development was seen when the Massachusetts Institute of Technology (MIT), in collaboration with the Hewlett Packard Corporation and the University of California, launched their own institutional repository systems, namely the DSpace software in November 2002 (Ware, 2004). This software is open source, and is being used by many universities to establish their institutional repositories. In January 2001, the University of Southampton in the United Kingdom also developed open source software for institutional repositories called ePrints. The development of this open source software greatly lowered the cost and barriers to implementing repositories for all institutions (Lynch, 2003). An increasing number of research universities around the world are now implementing institutional repositories.

2.2.1 The concept of institutional repositories

Institutional repositories are alternatives to traditional scholarly publishing. Institutional repositories are new models of storing and capturing scholarly output (Shearer, 2003). Broadly defined, a "digital institutional repository could be any collection of digital material hosted, owned or controlled, or disseminated by a college or university, irrespective of purpose or provenance" (Kim, 2007:1). An institutional repository is a digital archive of the intellectual product created by the faculty, research staff, and students of an institution and is accessible to end users both within and outside of the institution. Institutional repositories can be seen as vehicle for multi purposes, including the so called green route thus making published matreials available on an open platform. Institutional repository systems and services also signify new roles for librarians in an increasingly digital world as intimated by Walters (2007: 213) who stated that "when libraries create institutional repositories, they are reinventing themselves." Walters (2007) infers that as libraries establish institutional repositories, they become active producers, publishers and broadcasters. Libraries are therefore no longer passive receivers of information but active disseminators of intellectual output for their institutions. Libraries are in the process of redefining their roles, relationships and philosophical mindset (Walters, 2007).

2.2.2 Reasons for the emergence of institutional repositories

A number of scholars, namely Crow (2002); Lynch (2003); Prosser (2003); Shearer (2003); and Van der Merwe (2008), have identified many reasons for the emergence of institutional repositories. The most important is the high costs of scholarly journals. Van der Merwe (2008) argues that the current international library financial environment, budgets for published journals are generally low. She states that this can be attributed to library budgets becoming increasingly limited while the costs of scientific and scholarly journals increase annually by an estimated 15%. Institutional repositories, by providing free access to scholarly output, are a workable solution and alternative to the ever-increasing costs of scholarly publications (Van der Merwe, 2008). Bailey (2006) conducted a survey and discovered that the top three reasons given for implementing an institutional repository were "to increase global visibility, preservation, and the free provision of the institution's scholarship" (2006:14).

Numerous additional reasons for setting up institutional repositories have been proposed by different researchers. The institution, the individual, the researcher and the academic library involved in institutional repository initiatives recognise the benefits. For academics, the benefits include visibility and exposure of their research, resulting in increased citation of such works. As they are cited, they begin to be recognised and, over time their scholarly reputation increases. Another important benefit for authors is stewardship. This means that as authors write articles and publish them on open access repositories, their works are being preserved and managed in digital form. They no longer need to be concerned with maintaining and keeping the works safe in their computers or websites (Cullen & Chawner, 2010). Furthermore, for the institution, a published collection of journal articles through its open access institutional repository, provides the institution's research programme worldwide visibility and increases its impact (Chan & Swan, 2003).

Finally, for the individual, institutional repositories enable one to freely access, download and use research materials, which had not been previously possible. Although there is a rise in the establishment of institutional repositories all over the world, Shearer (2003) suggests that the success of institutional repositories will be determined by "their uptake, commitment and use by researchers" (2003: 106). She argues that the success of an institutional repository should be

determined by its use. One of the measures of usefulness of an institutional repository is contribution of content. People who contribute towards the content of institutional repositories include faculty, students and staff in universities. Faculty members, namely lecturers and researchers, are considered the most important contributors of scholarly content (Shearer, 2003).

2.2.3 The nature of content that institutional repositories collect

According to the Scholarly Publishing and Academic Resources Coalition (SPARC) Position Paper by Crow (2002), the content of an institutional repository should be "institutionally defined; scholarly; cumulative and perpetual; and open and interoperable" (Crow, 2002:16).

Institutionally defined means that institutional repositories collect and store original research output that have been produced by its members, such as researchers, academic and non-academic staff and students (e.g. masters and doctoral students). Institutional repositories also include theses, dissertations and other research-related outputs such as presentations, book chapters, research papers, articles and images. The repository may indicate the institution's intellectual output and academic quality (Crow, 2002).

The second term, *scholarly content*, means that institutional repositories focus on scholarly materials. One will find materials such as pre-prints of research articles, peer reviewed articles, published articles, monographs, theses and dissertations, and gray literature. A growing number of institutional repositories also contain books or book chapters (Chan & Swan, 2003). The authors of the books normally write them for financial benefits, and as such, they may not be willing to make the whole book freely available in a repository. They may choose to show the abstract, summary and the table of contents. In cases like these, the authors deposit the book with its metadata, publisher details and price. The text would be hidden to viewers and the book would be counted as a contribution to the institution's research output since the metadata will be visible allowing the book to be located by search engines. Chan & Swan state that the "evidence is accumulating, however, to show that when the entire content of a book is visible in a repository, sales of the book frequently rise" (2003:2). They argue that it is similar to what Amazon has with its "Look inside" facility. The visibility in the repository raises the book's awareness and promotes it to an audience that is more likely to buy it should it be relevant to their work.

Thirdly, the phrase *cumulative and perpetual*, means that institutional repositories should preserve and ensure ongoing access to materials they contain on a long-term basis (Prosser, 2003). Crow states that "providing long-term access to digital objects in the repository requires considerable planning and resources commitment" (2002: 18). Materials in the institutional repository are not to be removed once deposited in the repository. This calls for institutional repository managers to produce criteria, policies and implement management rights systems for repositories (Crow, 2002). An institutional policy must ensure that there are little or no hindrances to accessing the contents found in the institutional repository. This means that there should be free access to the contents. This could be done by allowing users outside the research community access to finding and retrieving information from the repository without encountering any problems (Crow, 2002). Therefore, an institutional repository should provide access through easily accessible search engines such as Google Scholar.

Lastly, an institutional repository should be *interoperable* with other repositories by means of multiple search engines and other discovery tools. Interoperability can be achieved by building institutional repositories on common international technical standards. This would allow different end users to search different databases and still locate and retrieve material within a repository. Making a repository interoperable increases its visibility and increases the citation number of the material in a repository (Pappalardo & Fitzgerald, 2007).

2.2.4 Responsibility for depositing materials in the repository

The people who contribute materials to the repository include faculty, students, university staff, and faculty members such as lecturers and researchers (Kim, 2007). These individuals are considered the most important contributors of scholarly content. However, authors such as Young (2002); Chan (2004); Foster and Gibbons (2005); and Pelizzari (2005) have noted that it has been difficult to get faculty members to contribute or deposit their work in repositories. This is one of the challenges that institutional repositories face.

While library staff and university administration have taken institutional repositories seriously, they have received little attention from academic staff who stand to benefit from them. Research shows that academic staff are reluctant to deposit their research outputs in the institutional repositories and that further research should investigate factors that contribute to and foster such

a reluctance (Cullen and Chawner, 2010). Furthermore, such studies should take cognisance of the perceptions and attitudes towards institutional repositories. Such research is necessary as the establishment of a repository is a major endeavour for any institution. The commitment of all its stakeholders is crucial for its success. These stakeholders may include university management as they control the financial aspects of the establishment, library staff, academic and non-academic staff, and students. Some of the non-academic staff would be involved in the technical aspects. For example, the Information and Technology departments may be tasked with the responsibility of maintaining the repository while the academics and students would be required to produce the scholarly content of the repository (Cullen and Chawner, 2008a).

2.3 Exploration, review and analysis of related concepts

To understand the concepts of institutional repositories, the following section explains the terms that have been used throughout this study.

2.3.1 Digital libraries

There are numerous definitions that summarize the meaning of digital libraries, depending on the context in which the terms are used. Institutional repositories can also be called digital libraries as they also offer access to intellectual research output in the digital format. A more comprehensive definition was formulated by the Digital Library Federation as follows:

> "Digital libraries are organisations that provide the resources, including the specialised staff, to select, structure, offer intellectual access to, interpret, distribute, preserve the integrity of, and ensure the persistence over time of collections of digital work so that they are readily and economically available for use by a defined community or set of communities" (Greenstein, 2000).

The emphasis in the above definition falls on the information being managed and organised in a systematic way and preserved for long-term basis. This definition means that digital libraries manage collections of digital objects that are systematically organised. The availability of digital objects are consistently supported by services that ensure their long-term accessibility and

quality. This enables users to retrieve them and exploit the resources through the use of digital equipment, namely computers that are connected to the internet.

2.3.2 Digital Information

Digital information may be kept in electronic format or created directly in digital form, (Smith, 2003) or materials that have been converted from hardcopy to computer-based materials through processes such as scanning. Digital materials range from simple text files to audiovisual materials. Examples of digital materials are word-processed files, files which are a combination of sound and images, and hyper links of digitised materials with other resources. Digital materials have become highly sophisticated web-based resources, which exploit the benefits of the technology used.

According to Smith, (2003) there have been numerous projects to digitise collections of text and images. The primary motivation has been the utilisation of digital technology to improve access to these materials, which would otherwise require a visit to the holding institution. Institutional repositories fall within these projects that digitise collections of research output (Smith, 2003).

2.3.3 Digitisation

The concept of digitisation refers to the process of "converting any physical or analogue item into a digital representation or facsimile" (Deegan & Tanner, 2002: 33). For example scanning a photograph and then saving the scanned image in a computer or flash drive. In so doing, one has converted a physical photograph into a digital presentation. Digitisation is most commonly associated with the preservation of cultural heritage materials via digital conversion in the interest of "providing the user unfettered access to a digital surrogate while also preserving the original" (Peters, 2003: 81). Cultural heritage institutions are not the only institutions engaged in the practice of digitisation. Most academic libraries store various kinds of information resources in a wide range of formats that are conducive to digital conversion, namely individual documents, bound print and manuscripts volumes, photographs, microforms and microfiche. The content of institutional repositories will include and therefore make digitised theses and dissertations, and data sets, available online.

2.3.4 Curation

Curation can be defined as

> "the activity of managing and promoting the use of data from its point of creation, to ensure it is fit for contemporary purpose and available for discovery and re-use" (Brophy & Frey, 2006: 15).

Brophy & Frey (2006) further state that curation is ensuring the long-term accessibility and re-usability of digital information. For datasets that keep changing it is also important to continuously update them and keep them fit for use. Maintaining annotations and links need to be done for higher levels of curation (Brophy & Frey, 2006).

2.3.5 Archiving

Archiving can be defined as

> "a curation activity which ensures that data is properly selected, stored, can be accessed and that its logical and physical integrity is maintained over time, including security and authenticity" (Hitchcock, et al., 2005:4).

Many research institutions thrive on having their research output accessed by users and researchers from other institutions. These research materials have to be maintained and easily accessible. Archiving is defined as a process of identifying, collecting and making accessible materials in their current format for immediate user retrieval by any archival institution or any function within an organisation (Smith, 2003; Brophy & Frey, 2006 and Hitchcock, et al., 2005). Some regard the process of preservation as synonymous with archiving. However, Phillips (1999) points out that while preservation is more concerned with the management of materials for long-term accessibility, archiving is more concerned with accessibility for immediate use. Institutional repository models "provide a means for institutions to create archives and make available their wealth of knowledge" (Prosser, 2003:168). It allows scientists and researchers to self-archive their own material. Institutional repositories therefore aid researchers and authors by archiving their work for future use.

2.3.6 Digital electronic archiving

Hodge and Carroll define digital electronic archiving (DEA) as the

> "as the long-term preservation of information published ("born digital") or communicated initially in electronic form (and perhaps in print as well). It does not include projects that simply convert legacy print information into electronic form for preservation and archiving. However, DEA may apply if the resulting electronic version is considered to be the primary or sole archive." (Hodge and Carroll, 1999: 17).

The above definition refers to the long-term storage, preservation and access to born digital information. It is concerned with *only born digital materials* which differs from traditional archiving which focuses on the accessibility of materials in their current hard copy format for immediate use. As defined above, **archiving** is the process of identifying, collecting and making accessible materials in their current format for immediate user retrieval by the archival institution or any archival function within an organisation.

2.3.7 Preservation

Preservation can be defined as

> "an activity within archiving in which specific items of data are maintained over time so that they can still be accessed and understood through changes in technology." (Hitchcock, et al., 2005:4).

This activity is important in ensuring that data is maintained in its original state for future use.

2.3.8 Interoperability

Interoperability refers to

> "the capability of a computer hardware or software system to communicate and work effectively with another system in the exchange of data, usually a system of a different type, designed and produced by a different vendor" (Reitz, 2006).

Hitchcock and others (2007) state that Open Archive Initiative (OAI), the Protocol for Metadata Harvesting (OAI-PMH) and the Open Archival Information System (OAIS) Reference Model

enables compliant sites to be *interoperable*, allowing both institutional and disciplinary repositories to be visible and viable. Interoperability has been discussed in detail in chapter 2 (cf.2.12).

2.3.9 Open access institutional repository

Crow (2002) provides a working definition of *open access institutional repository*, which is

> "a digital archive of the intellectual product created by the faculty, research staff, and students of an institution and accessible to end users both within and outside of the institution, with few if any barriers to access" (Crow, 2002: 16).

The development of open access institutional repositories was successful in the developed countries. Christian (2008) noted that there are some developing countries that have also done well in this aspect and gave an example of such countries. Among them are Brazil, India and South Africa. In South Africa, twenty-four academic institutions have open access institutional repositories. UP is considered the leading example as in 2010 was ranked number is 65 in the world and this was number one in Africa terms of content, size and visibility by ROAR (Hammes, 2010).

2.3.10 Data and metadata

Data and metadata are terms generally used to describe information content in digital libraries. Data can be described as information that is encoded in digital or computerized form (Arms, 2000: 14). Metadata is essentially data about specific data, which is generally divided into the following three groupings:

a) "descriptive metadata such as bibliographic information of a catalogued item,
b) structural metadata about formats and structures, and
c) administrative metadata, which is used to manage information" (Arms, 2000: 279).

Metadata that is stable and standardized makes certain that digital materials are available, and can be seen and used all the time Pasqui (2007). This means that when repository managers plan,

must also look into metadata preservation. Before any research material is deposited in an institutional repository, metadata should be added for the easy retrieval of the documents.

2.4 Digital Preservation

Digital preservation refers to the "series of managed activities necessary to ensure continued access to digital materials for as long as necessary" (Jones & Beagrie, 2002 as quoted by Hockx-Yu, 2006: 234). A number of factors have contributed to the production of large amounts of "born digital" information. These include the increase in computer power, network bandwidth, scientific instruments and gadgets. This data deluge has forced many researchers to address the issue of long-term preservation and curation as it means ensuring that data is not lost to the obsolescence of technology (Hockx-Yu, 2006).

Conway (1990) briefly defines the term preservation as the ability "to ensure protection of enduring value for access by present and future generations" (1990: 206). Preservation is best known in terms of protection of traditional information bearing formats such as paper, artefacts and analogue tapes. Cultural institutions such as libraries, archives and museums are widely recognised as the custodians of society's collective memory by preserving and providing access to the cultural, historical and intellectual resources of a society (Smith, 2003: 4). However, in today's globally digital world, all kinds of resources are increasingly being created, produced, distributed and accessed in the digital form. Smith (2003: 4) states that, "over 93% of new information produced is created directly in digital form" or is "born digital" (2003: 4). It is clear that digital information has become a vital resource contributing to our intellectual heritage.

Preservation of digital information poses challenges and problems of a different nature to the traditional preservation methods of information professionals. These challenges range from technical, social and cultural problems where cultural records that increasingly exist in digital form will require information professionals to focus on preserving the information for future access.

2.4.1 Digital preservation challenge

The most important things to consider when an organisation wants to deal with the digital preservation challenges is to first identify what is it that it wants to preserve. Secondly, the organisation should identify the users for whom they are preserving the information, then ask themselves what is the purpose of preserving such information is and finally ask themselves the time that will be taken to preserve the information they are interested in (Smith, 2004). From the literature as noted by Smith (2004) and Bellinger et al. (2002), preservation of digital objects involves a number of problems which includes, the problems of digital content, organisational problems (institutional roles and relationships), legal and regulatory problems (intellectual property rights), and the problems of the user.

The first challenge of digital content is that digital data is fragile because it is both intangible and transient. It also has a short life span as opposed to print documents, which are tangible and have a physical form. The challenges of preserving digital information stem wholly from its fragile nature since "it is malleable, unfixable, immaterial and very difficult to preserve" (Smith, 2004: 107). The fragile state of digital data dictates that the preservation of digital information should require more attention than that of other types of media. The physical form of other media help them to last for over hundreds of years, whereas digital information or data in digital formats may last ten years at most as the format changes quickly due to advances in technology. Open source formats however, will last longer. For example, fifteen years ago people used to store their data in floppy disks, which have been replaced by flash drives. Although some computers still have the disk drives, the information is not readable because it is not compatible with the latest software. The challenge with digital information is that it easily changes and should be updated from time to time

The second problem of digital preservation is that organisations such as libraries do not own electronic information and as such have no control over it. In that case the problem is that libraries cannot preserve information which they do not own. This then brings in the third problem of intellectual property rights. When it comes to digital information, copyright owners who are the people who produce or create the information might be ignorant about issues of preservation for their materials. Smith (2004) argues that preservation should then be a concern for both libraries and publishers and they should work together to do it. Lastly, users also

contribute to the problem of digital preservation as they publish on the web without considering any curation and long-term preservation plans of their materials (Smith, 2004)

Institutions therefore, need to spend a lot of money and resources to ensure the preservation of their digital data. According to McGill (2011) over the last decade many digital electronic equipment used to produce large amounts of digital data and information have been invented, and have caused a data or information explosion. Examples of these instruments include digital cameras which produce digital images and photographs, digital scanners used to convert hardcopy print materials into digital documents, facsimile machines and printers, cell phones, computers with sophisticated software that are compatible with the instruments, DVD and CD ROM writers, and digital recorders that produce music. Digital information created only a decade ago is in serious danger of being lost, thus creating a *'digital dark age'* (McGill, 2011: 18). Not only are librarians faced with the laborious task of preserving massive amounts of digital information, they also have to contend with the various challenges of preserving digital data (Smith, 2004).

The challenge for digital preservation is not just about the volume of data, it is also about the hardware and software used to store and access digital information (Hockx-Yu, 2006). Software and hardware are constantly upgraded and superseded. The debate for digital preservation continues, though currently no clear answers for institutional repositories have been found. According to McGill (2011), the Online Computer Library Centre (OCLC) developed a four-point strategy for long-term preservation in 2006 and proposed an OCLC Digital Archive Preservation Policy and Supporting Documentation. All institutions and organizations which are interested in formulating preservation policies can consult this document for guidance.

Another solution that can address the problem of digital preservation is that of institutions establishing institutional repositories in order to capture digital information created in their universities or campus.

2.4.2 Technology Obsolescence

It is important for institutional repository managers and implementers to have a preservation plan when establishing institutional repositories. They must think about the long-term future of the materials entrusted in their repositories. The term *obsolescence* means "a loss in the utility of an

asset due to the development of improved or superior equipment, but not due to physical deterioration" (Reitz, 2006). When used in relation to technology, it means that technology changes quickly. Institutional repositories must therefore play an important role in digital preservation and the long-term curation of digital materials. Technology change is estimated between two to five years and the original hardware and software becomes obsolete so institutional repository managers must ensure that file formats continue to be readable (Smith, 2003). Technology obsolescence is regarded as the greatest technical threat to ensuring continued access to digital material. This has been emphasized by Smith (2003) who stated that "digital information is highly susceptible to technological obsolescence and physical deterioration hence it requires coherent preservation strategies" (Smith, 2003: 5). Two methods that have been adopted by archivists to preserve digital information are called "refreshing and migration." Refreshing is method whereby archivists copy the data from old media onto new media to overcome the fragility of impermanence of digital media (Garett and Waters, 1996). This method is dependent on the compatibility of software and hardware versions from the past to the present (Garett and Waters, 1996). Migration is the periodic transfer of digital materials from one hardware/software configuration to another (Garett and Waters, 1996).

Although digitisation can be used as a method of preserving the original archival record, it is not a primary preservation medium (Dunlop and Hart, 2005). In other words, the digital image is used to provide access to the archival record, thereby avoiding excessive handling of the original, in that way the original is preserved as the digitised version will be used. This rationale was applied in the Bleek digitisation project at the UCT Libraries (Dunlop and Hart, 2005).

2.4.3 Preservation Policy

Various authors such as Smith (2003); Wheatley (2004); Hockx-Yu (2006); and McGill (2011) have proposed that institutional repository managers should prepare preservation policies for institutional repositories which answer the following questions:

- "For how long will the repository aim to preserve deposits?
- Can this be guaranteed?
- What formats should be used for preservation purposes?
- How is the repository intending to ensure continued readability and usability of its items?

- How are files in the repository backed up, in what form, and how often?
- How will content be updated, particularly when it becomes obsolete or wrong" (McGill, 2011: 34)

If repository implementers could plan for preservation, they can ensure that their repositories will serve its users for a long time.

2.4.4 The Open Archival Information System (OAIS) model of Preservation

The Reference Model for an Open Archival Information System (OAIS) of preservation is used for institutional repositories. It aims to describe the functions, roles and responsibilities of archival repositories. It also aims to establish a minimum level of functionality for archival repositories. This model provides a complete functional and information specification of a repository and "establishes mandatory responsibilities that an organization must discharge in order to operate an OAIS archive" (Hockx-Yu, 2006: 238). The functions required for digital preservation includes ingest, storage, retrieval, and long-term preservation of digital objects.

The OAIS Reference Model defines an OAIS as "an archive, consisting of an organization of people and systems that has accepted the responsibility to preserve information and make it available for a Designated Community" (Hockx-Yu, 2006: 239). In order to establish trusted repositories and also to ensure long term preservation of digital assets Hockx-Yu (2006: 239) asserts that repository managers should apply the OAIS principles which focus on the "implementation of Open Access archives in accordance with OAIS functional and information models" (Hockx-Yu, 2006: 239). Institutional repositories use this model in order to preserve the digital objects they contain.

2.5 Academic libraries as centres for research support

Libraries have a major role to play in the support of learning and teaching, research strategy, digital scholarship, and e-learning. Nfila (2007) asserts that the library should be seen as an important stakeholder when it comes to formulating policies that relates to teaching, research and learning. The UCT Strategic Plan and Priorities for the years 2010-2014 states that the library intends to "restructure library services to enhance support for meeting the information needs of

academics and postgraduates, while providing specialized, targeted services for undergraduates" (Rapp, 2009: 6). In addition, the library intends to promote and support the information needs of users through the development and establishment of the upcoming institutional repository.

From a long time, libraries have always been used by academic and research institutions for preserving and disseminating knowledge (Christian, 2008). This is the case since academic libraries have been responsible for storing books and other materials such as research outputs in the form of printed scholarly journals. Scholarly publishing disseminated research output in scholarly journals which were stored by libraries. The high costs of journals, together with the costs incurred by technological changes, has influenced the manner in which research output is disseminated. According to Cullen and Chawner (2009), for the past 150 years academic libraries played an important role in the scholarly communication cycle. Through the years, libraries changed the way they create, disseminate, and preserve knowledge. These changes were brought about by social, philosophical, and educational reform as well as the emergence of new technologies. Libraries have been able to adjust to and adopt a new technology, namely the development of institutional repositories that brings about changes in scholarly communication (Cullen and Chawner, 2009). This change resulted in the "networked information economy." It makes use of the internet through open access models of publishing done through the establishment of institutional repositories by libraries. This has gradually displaced the "industrial information economy" (Christian, 2008: 5). Academic libraries have been farsighted in advancing research, scholarship and knowledge but are now seen as leading custodians of research. They have to take up the leadership role in the establishment of institutional repository.

Authors such as Crow (2002); Lynch (2003); Chan (2004); Gibbons (2004); and Walters (2007) have stated that libraries are properly positioned to manage institutional repository development and sustainability, because of their existing information infrastructure and professional expertise. Crow puts it in this way:

> "In the long-term, organizing and maintaining digital content— as well as supporting faculty as information contributors and end users— should remain the responsibility of the library. Libraries are best suited to provide much of the document preparation expertise (document format control, archival standards, etc.) to help authors contribute their research

> to the institution's repository. Similarly, libraries can most effectively provide much of the expertise in terms of metadata tagging, authority controls, and the other content management requirements that increase access to, and the usability of, the data itself" (Crow, 2002: 20).

Libraries have always had the responsibility of taking care of the needs of their user, being able to guide them even when faced with technological and information overload changes. It is not surprising that UCT is now taking the responsibility further, by providing their users with online technologies and by developing and implementing an institutional repository. Furthermore, Amaral (2008) affirms that it is indeed the duty and responsibility of libraries to establish institutional repositories. It is an important and basic characteristic of libraries to collect and preserve information, which includes both the paper and digital output of an institution. In so doing, libraries affirm their role to "provide long-term access, open access and the improved, re-use of intellectual property" (Amaral, 2008:5). These activities correlate with the mission and goals of libraries.

2.5.1 The role played by libraries in preservation, curation and archiving

Preservation is associated with the protection of traditional information bearing formats such as paper, artefacts, and analogue recordings. Archiving refers to long-term storage or backup while an archiving system saves records for future use. Cultural institutions such as libraries, archives and museums have established preservation programs for traditional materials that they hold. The reason behind this is that they are recognised as the custodians of society's collective memory as mentioned earlier and they preserve and provide access to "cultural, historical and intellectual resources of society" (Smith, 2003: 4). This includes the regular allocation of resources for preservation, preventative measures to arrest the deterioration of materials, remedial measures to restore the usability of selected materials, and the incorporation of preservation needs and requirements into overall program planning. Today, materials such as electronic government documents, electronic journals, institutional documents and electronic theses and dissertations need to be preserved and archived. Furthermore, librarians are required to oversee the process of preservation and archiving. Libraries are therefore faced with the challenge to ensure that such information is accessible for future generations by means of

appropriate digital methods and strategies. Libraries, through the establishment of institutional repositories, are thus tasked with preserving scholarly content for long-term use.

2.6 Scholarly Communication

Libraries are no longer seen as storehouses of print materials because they now provide their users with quick information retrieval services and full-text electronic materials such as documents, reports, articles and journals. They are also expanding their services by collecting content in the digital form and providing contents to their users by digitising archival, special collections, images and theses and dissertations through setting up institutional repositories (Campbell-Meier, 2008).

Developments in information and communications technologies have resulted in a trend towards open access institutional repositories, which are digital archives for the preservation and dissemination of institutional research outputs. According to authors such as Prosser (2003); De Beer (2005); Campbell-Meier (2008) and Rieger (2008), scholarly communication has changed due to the widespread adoption of the internet. The internet has resulted in many improvements in scholarly communications since the 1990's and electronic publishing is seen as providing "greater diversity and choice" in scholarly communication and publishing (Campbell-Meier, 2008: 1). This has brought about changes in the way scholars and libraries publish, access and store their scholarly works in digital formats (Campbell-Meier, 2008).

The research library community introduced institutional repository systems in the early 2000's as a way to introduce a new scholarly communication model that built on new networking and storage capabilities in response to the changes made by the networked environment and the publishing industry (Rieger, 2008). In order to manage and preserve the digital content which universities are producing, academic libraries began to develop institutional repositories to house scholarly information. These repositories include traditional forms of scholarly communication such as journal articles and technical reports. The Carnegie Mellon University website provides the following working definition of scholarly communication:

> "Traditionally the term *scholarly communication* was narrowly defined as the system for disseminating scholarly work, primarily through journals. More recently the definition has been broadened to include the creation, transformation, dissemination and preservation of knowledge. It encompasses the entire process by which academics, scholars and researchers share and publish their findings within and beyond the academic community, and the entire gamut of publication types, from traditional journal articles, books and conference papers to sound and video recordings and interactive multimedia." (Carnegie Mellon University, 2011)

Scholarly communication comes in different ways such as telephone, traditional mail, e-mail, wikis, blogs, attendance at conferences, and formal peer-review (De Beer, 2005). The examples given above are of informal communication, which might also be form of scholarly communication. Scholarly communication is therefore not only limited to the act of publication, it includes social processes which are associated with research and publication (De Beer, 2005). For example, De Beer (2005) argues that teaching can also be viewed as an expression of scholarly communication since the learning environment is the one in which the next generation of scholars are introduced to the research domain. The library and information science community established and developed institutional repositories based on the premise that since the methods of scholarly dissemination were relying on commercial publishers there was need for change. Therefore, a change was effected which moved away from depending on commercial publisher to a system where institutions established open access institutional repositories thus the open access model of publishing (Rieger, 2008).

2.6.1 Stages of Scholarly Communication

The revolution of scholarly communication started with handwritten manuscripts. According to Alvi (2004), the 15th century, saw the coming of the Gutenberg Revolution which introduced the printed book. The coming of the Industrial Revolution later followed with its new developments and discovered that the book reached the people at a slow rate as books took a long time to be written and published. Not many people were able to get information from books quickly and thus delaying the disseminated of information (Alvi, 2004). Henry Oldenburg introduced the journal, which tried to address the problems that the book could not. This is the problem of

speed. The journal was smaller in size, having many pages and faster to produce and disseminate. With time, the print journals became expensive and different stakeholders such as libraries, authors, researchers, and readers came together to think of ways to deal with the traditional, subscription based business model for scholarly communication where journal prices had increased tremendously. This is commonly referred to as the 'Serial Crisis'. The crisis included problems such as:

- "too many journals were available,
- the time delay when journal articles came for submission and publication as well as refereeing abuses and problems of plagiarism, and
- the commercialising of periodical publications resulted in the inflated prices of journals as well as price discrimination and price disparities" (Alvi, 2004: 253).

The above problems prompted debaters to think of other ways to publish journals. Online communication technologies, namely the internet brought a solution to the problem of high subscription fees of journals. It brought with it the open access model of publishing which advocates for the establishment of institutional repositories. The internet has increased the pace of scholarly communication, "permitting a worldwide and almost up-to-the-minute transmission of scholarly ideas and information" (Johnson & Magusin, 2005:23).

2.6.2 Electronic publishing and publication

Electronic publishing is one of the ways of publishing scholarly communication. Over the last 15 years, worldwide access to the internet has seriously changed patterns of communication, both formally and informally (Pickton and Mcknight, 2007). Now scholars have more options to disseminate their work through the internet, for example, by choosing to publish their work on their personal, subject-based websites, departmental websites or in institutional repositories.

An *electronic publication* is "a document distributed primarily through electronic media" (Kling and McKim, 1999: 891). Examples of such publications include articles posted on websites, articles sent through email, or emails, electronic journals, magazines and newspapers, electronic books, weblogs and webpages. Some of the electronic publications are online versions of print publications (Reitz, 2006).

In the process of harnessing the power of information and computer technologies, the internet is competing with print publications but it may also be regarded as a substitute of print publications (Alvi, 2004). Although the journal is still the most important form of disseminating scholarly communication, it is now available in different formats as apposed to traditional print format.

2.7 Theoretical Framework for Institutional Repositories

Institutional Repositories are derived from the conceptual models supporting electronic scholarly communication forums (eSCFs) and digital libraries as well as knowledge management systems (cf. 1.2). In this study, institutional repositories will be discussed as a new strategy for facilitating changes in electronic scholarly communications through the open access model of publication (Crow, 2002; Lynch, 2003; Shearer, 2003; and Chan, 2004). The following section discusses the open access model of publication on which institutional repositories are based.

2.7.1 Exploring the Open Access Model of publication

The Budapest Initiative (2003) defines open access to literature by stating that:

> "its free availability on the public internet, permitting any user to read, download, copy, distribute, print, search or link to the full texts of these articles, crawl them for indexing, pass them as data to software, or use them for any other lawful purpose, without financial, legal, or technical barriers other than those inseparable from gaining access to the internet itself." (BOAI, 2002).

The above definition implies that all readers and researchers who have access to the internet are freely able to read, download and use the texts they find. The concept implies that open access is freely available through the internet, that there is a peer-review process, that the author holds copyright and that the content has an academic purpose. When making an article available through open access, the publisher does not charge the reader to access the article and also does not restrict the copying or redistribution of the article in any way. The concept of open access is based on the premise that academics are interested in the distribution of their work as widely as possible without financial barriers, rather than wanting to profit from their knowledge production (Park & Qin, 2007).

Although the open access movement advocates for the free access to information, there are some hidden costs to the availability of the open access publications. The reader must first have access to connectivity, content and usability (Borgman, 1999). Connectivity refers to having the physical connection to use a computer network and not the free connectivity of physical networks. Content means that the information must be present or exist in the computer networks, and usability means that one is able to use the computer networks and the content by being able to search, read and disseminate information.

2.7.2 Policy and Principle statements about open access

Open access started to receive attention from organisations around the late 1990s. These organisations produced policy and principle statements regarding open access (Pappalardo and Fitzgerald, 2007). The statements were giving researchers and research communities around the world guidelines and explaining how open access can be done successfully. They further encouraged researchers to participate in open access. The statements also helped universities and research institutions to formulate their own institutional open access policies and principles.

The following discussion summaries the primary international open access statements:

2.7.3 Bermuda Principles - 1996

The Bermuda Principles are one of the first international statements on open access, published in 1996 (Pappalardo and Fitzgerald, 2007). Scientists who wanted to have fast and free sharing of their pre-published research notes on gene sequences, developed these principles. They wanted their research to be freely available on the web so that the research community could benefit from it (Pappalardo and Fitzgerald, 2007). The Bermuda Principles were approved in January 2003.

2.7.4 Budapest Open Access Initiative - 2002

Developments in open access continued in 2002, when the Budapest Open Access Initiative (BOAI) extended open access principles to peer-reviewed journal literature (Pappalardo and Fitzgerald, 2007). This organisation was developed in 2001 at a meeting of the Open Society Institute in Hungary. The aim was to remove access barriers to scientific literature. This would

be achieved through encouraging the self-archiving of published articles and open access journals. BOAI's open access statement reads thus:

> Removing access barriers...will accelerate research, enrich education, share learning of the rich with the poor and the poor with the rich, make this literature as useful as it can be, and lay the foundation for uniting humanity in a common intellectual conversation and quest for knowledge (Pappalardo and Fitzgerald, 2007: 3).

By November 2011, 6132 individuals and 597 organizations had joined the BOAI (BOAI, 2011 http://www.soros.org/openaccess/list_signatures).

2.7.5 Bethesda Statement on Open Access Publishing – 2003

After the publication of the BOAI open access initiative statement, many researchers as well as organisations worldwide started to support the open access movement. This led to more progressive statements on open access, which was taken up by academic institutions, international organisations and funding bodies. These include the Bethesda Statement on open access publishing and the Berlin Declaration, as noted below:

2.7.5.1 Bethesda Statement

The statement explains the concept of an Open Access Publication, as follows:

> "The author(s) and copyright holder(s) grant(s) to all users a free, irrevocable, worldwide, perpetual right of access to, and a license to copy, use, distribute, transmit and display the work publicly and to make and distribute derivative works, in any digital medium for any responsible purpose, subject to proper attribution of authorship, as well as the right to make small numbers of printed copies for their personal use.
>
> A complete version of the work and all supplemental materials, including a copy of the permission as stated above, in a suitable standard electronic format is deposited immediately upon initial publication in at least one online repository that is supported by an academic institution, scholarly society, government agency, or other well-established organization that seeks to enable open access, unrestricted distribution, interoperability, and long-term archiving (for the biomedical sciences, PubMed Central is such a repository" (Pappalardo and Fitzgerald, 2007:4).

The above quote provides a detailed explanation of what open access means and the parameters within which it operates.

2.7.5.2 Berlin Declaration – 2003

This Berlin Declaration on open Access to Knowledge in the Humanities, was published in 2003 in order to raise awareness of and support for making resources available through the internet and allowing access through online repositories. It also provides the definition of Open Access which resembles that given by the Bethesda Statement. The Berlin Declaration has been significant in advancing open access to research and scholarly material and by March 2007 had been signed by 227 organisations from all over the world. By November 2011, 6132 individuals and 597 organizations had joined the BOAI (BOAI, 2011). During the open access week, which was celebrated by UCT Libraries from the 23rd – 26th of October 2011, the UCT Vice Chancellor, Dr Max Price signed the Berlin Declaration on Open Access to Scientific Knowledge on behalf of the university. This marked an important step by UCT to show both its support for open access practices and showing its commitment to increasing access of its research output to Africa and the world.

2.8 Degrees of Open Access

The Open Access research literature extends beyond journal articles. According to De Beer (2005: 8), open access can be expressed primarily through four avenues:

- "Publication in Open Access journals;
- Making research available in an institutional or disciplinary (subject-based or topic) digital archive/repository;
- Making research available via Departmental or Personal homepages;
- Making the research output of postgraduates available via Electronic Theses and Dissertations (ETD) digital repositories".

2.8.1 Routes to the publication in Open Access Journals

Prosser of SPARC Europe defines open access journals as those which...

> "...do not charge for access to the papers, but make the papers available to all electronically and look to other financial models to cover the costs of peer-review and publishing. They do not invoke copyright or exclusive licenses to restrict access to the papers published within them, rather they encourage the dissemination of research limited only by the reach and extent of the internet" (Prosser, 2005a: 12).

A number of universities now provide support and funding to cover open access publication by their researchers. There are two ways of using open access to increase the reach and exposure of journal articles as noted below.

2.8.1.1 Green Route

The Green Open Access route refers to the process whereby an author places a copy, of the scholarly output, known as self-archiving in one or more Open Access repositories. These may be an institutional repository, a subject-based repository, or a combination of both. The business model for green Open Access publishing is that the body maintaining the repository pays for the ingest of materials, the addition of metadata and other technical and administrative requirements (Oppenheim, 2008). The advantages offered include immediate and wider circulation of the articles concerned and access to the literature by scholars, which could not normally afford journal subscriptions and readers from the university community (Oppenheim, 2008).

2.8.1.2 Gold Route

The Gold Open Access approach involves the publication of peer reviewed articles in open access journals that do not charge subscriptions fees to articles. They can be in electronic or parallel published journals. The business model is called 'author pays,' which according to Oppenheim (2008) is a misnomer because it is the author's employer or funder who pays, not the author him or herself. The cost of publication, instead of being recovered from subscribers, is paid by the publisher or the organization funding the research concerned, or from donor funding or scholarly society support or from a combination of all of these (Oppenheim, 2008: 580). Open

access journal publication offers a particular advantage to developing countries, thus offering greater citation impact and increased readership and reach (Oppenheim, 2008). Open access journals (gold route) can be seen as a more active alternative to traditional scholarly publishing.

2.8.1.3 Types of Open Access publications

Open access publications can be one of three types. The first is when an author submits a version of an article, which has been submitted but not yet accepted for official publication. This is known as a **pre-print.** The "term 'pre-print' is sometimes confusingly used to refer to draft versions of articles or manuscripts, regardless of whether they are destined for formal publication" (Kling, 2004: 598). When an author submits a peer-reviewed version of an article that has been accepted for publication, that is either waiting to be published or has been published, is known as a **post-print.** The post-print is the version of the article as it will appear in the journal. Lastly, the author may have an electronic version of a pre-print or post-print which is known as an **e-print** (De Beer, 2005: 9).

In South Africa, the gold route journal publications of local journals are supported by the Department of Science and Technology (DST), through the Academy of Science journal publishing initiative (Cox, 2010). These journals receive support from the DST for their publication on this online open access platform, where they benefit from the addition of meta-tagging for local and regional impact measurement (Cox, 2010).

2.8.2 Disciplinary digital repositories

These are digital collections of research articles in specific subject discipline repositories. Open access repositories may be centralized and subject-based, such as the long established physics archive, arXiv repository or CogPrints repository in cognitive sciences (Ware, 2004). According to Ball (2010), the first e-print repository, arXiv for physics was established at the Los Alamos National Laboratory in August of 1991. At first, it served only high-energy theoretical physics. The reason being that physicists had found out that journals disseminated information slowly as compared to when they sent each other completed hard-copy pre-prints of their articles. The physics arXiv repository later expanded to cover other areas in the science field, such as mathematics, computer science and even other areas in physics.

The physics arXiv repository success was followed by other disciplines and institutions also establishing repositories. This eventually led to the formation of the Open Archives Initiative (OAI) (Van de Sompel & Lagoze, 2000). In 2000, ePrint software which was adapted from CogPrints software was released after the first meeting of the OAI in order to help institutions to establish their own repositories. This was one of the developments from the first OAI meetings (Ball, 2010). Since then, other institutional repository systems have emerged, notably DSpace and Fedora.

2.8.3 Departmental repositories or personal homepages

Open access can also be expressed through the use of departmental repositories and personal websites, where researchers and students in a department or faculty may have a repository for submitting their research articles and papers. At UCT, departmental repositories, such as the UCT Computer Science Research Document Archive and the LawSpace from the Faculty of Law, are examples of departmental repositories which serve researchers and students and where student theses and dissertations are deposited.

2.8.4 Electronic Theses and Dissertations (ETD) digital repositories

Finally, open access can be expressed through a repository that focuses on electronic theses and dissertations. Open access institutional repositories are becoming one of the best avenues utilized by institutions in making their research knowledge or output widely available and accessible to the outside world via the World Wide Web (Christian, 2008).

From the literature, it is clear that many academic institutions carry out projects that establish electronic theses and dissertation repositories in order to showcase their research (Moahi, 2008, Christian, 2008). Electronic theses and dissertations repositories promote access and dissemination of scholarly research to a wide scholarly audience. Many universities are now requesting students to submit digital copies of their theses and dissertations for electronic deposit.

African universities are also welcoming the initiative of making their theses and dissertations accessible to researchers and the global knowledge pool, as in the past these remained inaccessible (Moahi, 2008, Christian, 2008). Institutions may establish institutional repositories

which are divided into two parts. Firstly, the institution may choose to establish a repository which focuses only on electronic theses and dissertations and later establish a repository which houses research output by academic staff and the institutions stakeholders. That is how the UP established its two repositories (Bothma et al., 2008). This research work will focus on both aspects of open access institutional repositories.

2.9 Factors for establishing a successful institutional repository

This will discuss the six factors that have been identified in the literature which are used to evaluate the success of an institutional repository. These factors are prerequisites for establishing an institutional repository.

a) Identify important role players,

b) Address issue of resources,

c) Evaluate software,

d) Formulate a policy for the institutional repository,

e) Restructure the library to address change, and

f) Licensing.

These six factors are discussed below and will be referred to in chapter 5 where they will be used for a comparative analysis between the UCT and UP institutional repositories.

2.9.1 Identify important role players

In order to establish an institutional repository, major role players have to be identified as they spearhead the project. According to Olivier & Louw (2011), the "role players in the ETD process are the academic community at large, funding agencies, the institution, the library, the supervisor, the student and the graduate school" (2011:1).

2.9.2 Address issues of resources

When an academic library plans to establish an institutional repository, issues of resources must be addressed. Questions such as whether the necessary resources for the institutional repository project are available, what the institutional repository will cost and possible sources of funding must be dealt with right from the start (Nabe, 2010). An assessment of resources is important in the planning stage of the institutional repository project, as the library has to seek support from relevant stakeholders. Nabe (2010) proposes carrying out surveys of what the library has, and of finding the cost of establishing the institutional repository in order to help the planners with the budget. Assessing the cost of an institutional repository primarily takes into account two main components, namely, the equipment (hardware and software) and the personnel required. Marketing and training may also need to be accounted for, based on which platform is chosen (Nabe, 2010).

2.9.3 Evaluate software

Both open source and proprietary platforms are available for institutional repositories. Open source software is free and proprietary software requires institutional subscription to a commercial provider. It is important for an institution to choose software based on its needs. It has been observed that "one of the key questions that determine platform selection is the availability of staffing" (Nabe, 2010: 29).

2.9.4 Formulate Policy for the Institutional Repositories

The need for a policy to guide the operation of an institutional repository is an important factor to be considered. Some issues include,

a) "what to accept or not to accept,
b) copyright issues,
c) self or mediated archiving,
d) submission and withdrawal policies,
e) types of materials to accept and

f) other issues governing the operation of the institutional repository form part of the policy of an institutional repository" (Asamoah-Hassan, 2010).

These aspects would need to be discussed and agreed upon by the institutional repository project committee. The policy should clearly indicate whether the repository is an open access initiative, and whether it would include abstracts as well as full text publications.

The German Initiative for Networked Information, also known as *Deutsche Initiative für Netzwerkinformation* (DINI), recommends that repositories should formulate policies and make them available to everyone, specifying how they will be run and what their content will be (DINI, 2011). The policy must also specify the rights and obligations of repository operators and of authors/editors. According to Corletey (2011), the policy should include a clear statement on the following areas:

- "An introductory statement on the objectives of the policy itself and indicate the scope of coverage,
- The objectives of the institutional repository and its benefits to the institution, the faculty and other staff, students, and the country,
- The scope of coverage should also be stated, i.e. the items to be held in the repository.
- A clear statement on who constitutes the stakeholders,
- The formation of a management committee to oversee the implementation of the policy (If possible the deputy should be chairing that body),
- A request for at least one IT support staff,
- A clear statement on the equipment required to run the platform. This should include at least two new servers, one high-end and a medium-end server. The mid-end server will provide back up if the main server fails. A book scanner and at least one computer,
- A clear statement indicating that the repository server should be provided with a public IP address and a dedicated bandwidth so that users can access it easily,
- A statement on the location of the repository server. But in all situations the access/control of the server should be under the Library,
- Statement on the standard format that all theses/dissertations should meet. This should include such things as form i.e. digital, pdf, arrangement of parts etc.

Theses/dissertations should have abstracts and keywords provided by the author of the work and approved by the supervisor before submission,

- Statement that students should submit their final and corrected theses to the library and be cleared by the library before graduation" (Corletey, 2011: 10).

The above guidelines should help the institutional repository managers to handle complaints.

2.9.5 Restructuring the library to accommodate change

Whenever there is change in any organization it is important for management to restructure and generate new staffing roles and responsibilities. In this case since libraries are moving towards establishing institutional repositories it will also involve restructuring to accommodate the implementation of a digital repository.

2.9.6 Licensing

In order to ensure open access to its collection, an institutional repository needs to be registered with open access harvesters such as OAISTER, Registry of Open Access Repositories (ROAR), Open DOAR and Google Scholar. This allows the content to be harvested by search engines for visibility. These search engines would then be able to pick contents from the institutional repository (Stevenson & Hodges, 2007). Licensing is what DINI refers to as *Visibility of the Service.* A repository should have an entire range of services, which are available via a website. The institutional repository's homepage must be accessible from a central location on the institution's homepage. According to the DINI criteria, an institutional repository has to be visible not only to the immediate and individual user, but also to external services such as search engines or other referencing services so that it can be visible to the whole world (2011). In addition to registering with the search engines it must also be registered and listed on the DINI website.

Although a repository can be harvested by providing OAI-PMH interface, it is still important to register repositories, as it is a useful means of promoting the visibility of repositories to service providers for harvesting. The Open Archives Initiative provides a

service allowing repositories to be registered as a data provider in the OAI registry (JISC, 2011).

2.10 Criteria for a successful institutional repository

The criteria for the success of an institutional repository would normally be taken from its statement of purpose. An institutional repository can only be successful if it fulfills its stated purpose which includes "preserving information and making it available for a designated community" (Thibodeau, 2007: 1). Another important factor to note when judging the success of an institutional repository its context, the purpose it serves and the environment in which it operates. Thibodeau (2007) is of the view that "the metrics for estimating success against the criteria must be formulated in light of the culture, constraints and opportunities existing in the environment" (2007:1). Nabe (2010) is of the view that institutional commitment is important for the success of an institutional repository. This is because it is an inclusive service which requires the involvement of everyone in the institution. He further argues that although libraries are the developers of institutional repository in academic institutions, they need the support of a number of departments in the institution. The technical services departmental staff needs to be involved with metadata, preservation standards and implementation. The IT department needs to be involved and provide their expertise, while other staff members may be involved in keeping the system running. The library administration may be reporting the institutional repository progress and advocating for its support from the University leadership. This is important as it shows that the institutional repository is one of the major efforts of the library to advertise the university to the world. If the institutional repository is successful, the university and its research output will be more visible.

Some of the criteria that determine the success of an institutional repository are discussed below. DINI and DRIVER (Digital Repository Infrastructure Vision for European Research) have noted additional criteria for success of institutional repositories. DINI suggested the following eight criteria:

a) "Visibility of the Service,

b) Policy,

c) Support of Authors and Publishers,

d) Legal Aspects,

e) Information Security,

f) Indexing and Interfaces,

g) Access Statistics, and

h) Long-Term Availability" (DINI, 2010:12).

DRIVER proposes guidelines which focus on the following five issues:

a) "collections,

b) metadata,

c) implementation of OAI-PMH,

d) best practices and

e) vocabularies and semantics" (2008:10-19).

DRIVER is concerned with textual resources that can be harvested with OAI-PMH (2008 a).This is because the major resources provided by digital repositories are text and the major approach for offering these textual resources is the Open-Archives-Initiative Protocol for Metadata-Harvesting.

The DINI Certificate provides a solid basis for what to consider when operating a repository. Since DRIVER looks at repositories from the perspective of an aggregator, the DRIVER Guidelines do not cover the aspects described in the DINI - Certificate which is designed for guiding the overall local operation of a repository (DRIVER, 2008b). Below is a discussion of some of the criteria that have been identified with which to determine success of an institutional repository. These criteria are drawn from both the DINI criteria and the DRIVER guidelines.

2.10.1 Content

The content of an institutional repository is the most important factor that shows the success of a repository (Organ and Deveson, 2010; Thomas, 2007; Harnad and McGovern, 2009). Content recruitment is key because it literally is the core of the institutional repository. Harnad and McGovern (2009) are of the view that not many researchers agree on what criteria to use to judge a successful repository. What seems to be clear is that content recruitment is an important factor to show whether the institutional repository is successful. A critical mass of material is necessary to generate both additional content and end-user activity. Value-added services including everything from full text retrieval to preservation are also seen as an important part of success (Harnad and McGovern, 2009).

The contents of an institutional repository should include both born-digital and the older repurposed digital materials. Westell argues that there should be more "born digital" materials in an institutional repository than repurposed materials as "the larger the critical mass of documents in an institutional repository, the more it will facilitate output measures" (2006: 216). This means that if a repository has a lot of content, it will stand a higher chance of being visible to the outside community. It will be used, and therefore, its citation rate should go up. *Long term availability* is a further criterion found in the DINI certificate and it can be discussed as part of the content factor. It concerns how the repository intends preserving contents for future use. DINI's minimum requirement is that a document should be archived in a repository for no less than five years.

2.10.2 Use

Use is the second factor that has been identified as a success factor and can be divided into three sub-categories: *number of users*, *type of content used* and *nature of use* (Harnad and McGovern, 2009). The most common method used to measure the use of a repository, is by webometrics. According to Organ (2006) and Thomas (2007), "webometrics is the statistical technique of counting the gross number of items in the live repository and to measure retrieval by counting hits and downloads with the option to sort by country" (Organ, 2006). Webometrics show how many hits have been made

from the repository and how many articles have been downloaded. The more the downloads, the clearer it becomes that the institutional repository is being used. To the researchers whose papers are downloaded, it may mean that they receive more citations. Westell (2006) noted that statistics of this sort contribute to institutional research outcomes and are powerful data to be used for career progress of individual authors.

As research outputs are increasingly exposed to the web via search engines, such as Google, organisations will increasingly be making use of institutional repository statistical packages. According to Organ (2006), high citation rates of researchers' work in various institutions make Vice Chancellors and CEOs accept these new kinds of research infrastructure, with the promise of improving and enhancing the reputation of their institution and research staff. In addition, bringing together input data, usage and citation analysis is helpful in providing the true picture of a repository's effectiveness. These data give evidence that the institutional repository is being used by a variety of users and also provides benchmarks of use and growth over time for institutional repository funders and university administrators (Westell, 2006).

According to DINI, the factor of use, which is measured by webometrics, is similar to its criterion of *Access Statistics,* which looks at access statistics of documents and publication repository as the qualitative, quantitative and technological basis for evaluation of the repository (DINI, 2011). Therefore, every repository must log statistical data. This can be achieved by a tracking system of downloads and citations. This is similar to bibliometrics and in websites it is known as webometrics.

2.10.3 Submissions

The rate and number of submissions which lead to growth of the repository are the third factor to consider in judging if the repository is a success. These are referred to as "repository deposit activity measures" (Thomas, 2007:138). This means that "sustainable repositories are only possible with sustainable deposit profiles through active community engagement" (Carr and Brody, 2007: 18). Institutional repositories should avoid the occasional high volume deposits. Carr and Brody (2007) proposed that "regular but less high volume inputs are a preferred way to sustain institutional

repositories, describing it as the difference between the "gappy" and continuous deposit profile (2007: 18)." Carr and Brody (2007) propose the use of the Registry Service ROAR (Registry of Open Access Repositories) to examine the performance of repositories according to the daily activity of any of its repositories. This factor of submission of deposits can also be sub-divided into the following categories:

- *Number of submissions* refers to the digital content that is deposited in the institutional repository by the person who produced it.
- *Frequency of submissions* refers to deposit being made continuously. As noted above it is recommended that submissions should be continuous over time and not just large deposits occasionally.
- *Type of submitter* refers to the people who deposit contents in the institutional repository. They range from staff, graduate students and undergraduate students. Researchers advocate for a broad representation of constituents in the institutional repository in a university and want all departments to submit content.
- *Participation of key stakeholders* refers to participants of university administrators and funders where they too submit contents in the institutional repository (Harnad and McGovern, 2009).

2.10.4 Advocacy

Advocacy is the fourth success factor to be considered. Advocacy deals with building informed awareness for which a core message and ethos are essential (Johnson, 2008). According to Pappalardo and Fitzgerald (2007), it is important for repository implementers to make sure that staff and students in the institution are informed and aware of the existence of the institutional repository. They must be fully informed of the reasons for establishing the institutional repository and how they can deposit materials. Institutional repository implementers should ensure that they provide open access policy to staff and students as they are the main depositors of content (Pappalardo and

Fitzgerald, 2007). In order to ensure that the message of Open Access and the benefits of establishing and managing institutional repositories is spread as effectively as possible, staff members need to create a communication plan for the advocacy campaign. A number of different advocacy strategies can be used, including top-down and bottom-up, blanket and targeted approaches (DRIVER, 2008b).

Advocacy strategies range from using informational websites, promotional brochures, leaflets, presentations to departments, faculty presentations about open access and the digital repository, individual meetings with academics to publishing articles in in-house newsletters, magazines and campus newspapers to inform staff of the progress the repository makes over time (Pappalardo and Fitzgerald, 2007). Westell (2006) calls this promotion and argues that a repository must be promoted to faculty staff in order to inform them of open access publishing, as it may be a new concept to them. Promotion is an important factor that can help populate the institutional repository once academics understand the changing culture of scholarly communications.

2.10.4.1 Addressing authors' concerns

Pappalardo and Fitzgerald (2007) stated that an essential part of institutional repository advocacy is to address authors' concerns about the repository and the deposit process. They noted that the most common concerns by authors fall into the following three main categories discussed below:

- *Concerns about the repository itself* – many academics are worried about the appearance of their articles after they have been deposited in a repository. They are unsure whether their articles would be formatted correctly and displayed at a high standard, and whether these articles would easily be retrievable through use of different access points such subject, author, title and year.

- *Concerns about their rights* – many authors are not sure whether their publishers will permit them to deposit their work on institutional repositories. They are unaware of the rights they possess under copyright law. Another concern is about

self-archiving. Most authors tend to think that self-archiving is the same as self-publishing. They are afraid that if they self-archive, publishers will not allow them to publish as their work will already be available through institutional repositories. Pappalardo & Fitzgerald (2007) state that this is a legitimate concern as some publishers will be extremely strict about the release of pre-published material. Therefore, it is important for institutional repository managers to explain to authors that materials can be published traditionally and also deposited in a digital repository after all the necessary permissions have been obtained from the publisher. Another related factor is that authors might think that their work would be plagiarised from a repository.

The advice given to institutional repository implementers by Pappalardo & Fitzgerald (2007) is that their open access policy and Repository Deposit License should clearly specify that depositing articles into an institutional repository does not transfer copyright of their materials. Institutional repository implementers should also advise authors to read their publishers' websites about the publishing agreements and publishers policies on open access.

- *Concerns about the deposit process* - institutional repository implementers should explain to academics and students what the deposit process entails and also clarify that it takes about six to ten minutes to deposit their first article, and only a few minutes once they are familiar with the process. This is because academics do not want to waste their time. They are mostly concerned about the time and effort they would use to deposit their work, entering metadata and the time to seek permission from the publisher to deposit the work (Pappalardo & Fitzgerald, 2007).

Pappalardo & Fitzgerald (2007) urge institutional repository implementers to inform academics and researchers of the following benefits of open access repositories which include:

- that the authors' work will be disseminated quickly;

- their work in the repository will be easily found through searching by using different search engines; and
- their visibility through downloads of their articles will boost their research profiles and the institutions' profile (2007:60).

It therefore calls for institutional repository implementers to motivate researchers and authors by showing them details about how many times their articles have been downloaded, cited and accessed. This can make authors enthusiastic about depositing their work in an institutional repository.

For an institutional repository to be successful, cultural change must be achieved through advocacy or "getting the right message to the right people with the tone and content varied by audience" (Johnson, 2007a: 23). Though the institutional repository may have funding, staffing, and fundamental values and policies, it still needs advocacy to inform stakeholders of their value and use.

2.10.4.2 Pillars of Advocacy

SHERPA observed that neglecting advocacy resulted in repository decline and stagnation but, with the right level of engagement across the institution, advocacy became the key to success. Advocacy was the main focus of the SHERPA Partner activity and the cornerstone of achieving cultural change within institutions echoes Johnson (2007c: 24-27).Some of their "pillars" include:

- *"Set achievable targets"*: When forming steering committees for advocacy campaigns bring in people who are enthusiastic about open access and who can drive the institutional repository developments. These people should be good at communication so that they could face and resolve issues quickly. It helps to focus on specific targets.
- *"Discipline and community"*: The differences between subject repositories and institutional repositories must be clarified so that academic staff may deposit their work in the institutional repository. This should be followed by encouraging academic staff from all disciplines, departments and faculties to deposit materials

in the institutional repository. The committee should also encourage long-term stability of the institutional repository. The committee may further provide support to the entire academic staff as a major advantage of depositing locally.

- *"Educate and clarify"*: It is important to educate and clarify issues to academics about the global nature of the challenges of scholarly publishing. This will assist researchers to be aware of challenges they may face if they want to publish in reputable journals. It will further assist researchers and authors to deal with the issues which might arise in publishing.

- *"Seize the moment"*: Marketing the repository is an important activity. It should be done whenever an opportunity arises in order to interest depositors in the institutional repository.

- *"Enable effective deposition"*: Effective deposition is the final advocacy goal. This can be done by institutional community awareness programmes. Advocates may approach publishers where OA rights information is unavailable, or search for new articles published by academics, and then seek permission to deposit from the academics.

- *"Achieve quick wins"*: Enhanced rapid population of a repository can be possible by advocating more personnel, provided that repository managers coordinate and ensure that all staff members are involved. These helpers need to be aware of planned future enhancements, as well as a core service goal. Involvement of administrative staff in departments as depositors on behalf of their academics has proved relatively successful, helping to overcome academic time-concerns over deposition. If properly briefed, administrative staff can act as effective advocates themselves. This is an especially powerful tool for those institutions with more limited staff resource (Johnson, 2008).

- *"Facing the challenges"*: There are many challenges for advocating repositories. They can be technological, lack of awareness, or time constraints that prevent authors from depositing their work. All of which frustrate administrators' efforts. Usability testing and effective customization of repository platforms could

address these challenges. This means having simple clear instructions and guidelines on the repository website on how to submit. This should be step-by-step instructions so that it will not waste the depositors' time (Johnson, 2008: 24).

2.10.5 Support

Support is the fifth factor to consider when judging the success of an institutional repository. Without support from the university community and the stakeholders, the institutional repository will never be successful. Support comes in three forms, which are:

- *Constituent support* – The community that the institutional repository serves should support it. Content depositors should be willing to deposit their work without being persuaded first. Users should express their satisfaction with the institutional repository by giving institutional repository feedback. For example, through blogs or emails.

- *Financial support* – It is important for the setting up and maintenance cost while the institutional repository is running. The institution can provide ongoing, and preferably increasing support or the institutional repository may have funding from outside international organizations.

- *Technical support* – When setting up repositories it requires technical help and support. Thus "the initial set up of an institutional repository requires a mixture of library and IT skills, and these are not always in place" (Markland and Brophy, 2005: 13). There should be adequate support for, and interest in the development and enhancement of the institutional repository infrastructure as well as, the necessary software and tools required (Harnad and McGovern, 2009). Staff needs to be trained in order for them to work effectively and efficiently.

Possibly the surest way to guarantee support for an institutional repository is to mandate participation which can only be done at university administration level (Nabe, 2010). This is said to be the most important role of the university administration, but not all

universities mandate open access for research by staff and students. Nabe (2010) argues that the need for support goes beyond the library which, as the developer of the institutional repository, it encompasses everyone in the university community.

2.10.6 Influence

Influence refers to providing assistance to other institutions in establishing their own institutional repositories. This is similar to what Thibodeau (2006) calls collaboration. It is about how the institutional repository associates with others in the country, region and in the world. Institutional repositories are encouraged to work in collaboration with others to demonstrate that they are successful. If an institutional repository works in isolation, it might signify failure because it does not share and transfer knowledge. A more collaborative arrangement would exist where separate institutions decide to work together to leverage each other's strengths and where they form or join consortia for such purposes (Thibodeau, 2006).

2.10.7 Interoperability

Lastly, an institutional repository should be interoperable. This refers to the capability of a computer hardware or software system to communicate and work effectively with other systems in the exchange of data, usually a system of a different type, designed and produced by a different vendor (Reitz, 2006).

The importance of interoperability is that OAI and other search engines, such as Google Scholar, can harvest their contents. It is therefore advantageous to producers of repository contents as it exposes scholarship in a new way and puts it in an international context. The repository may be the only vehicle to distribute this material (Westell, 2006). A major success factor for an institutional repository is to meet the universal standards of OAI-PMH as it increases the accessibility of its content (Carpenter, 2008). There two approaches that are used to achieve interoperability; these are cross-system searching and metadata harvesting. A protocol that is used for cross-system searching is the Z39.50 protocol while with harvesting OAI uses OAI-PMH is used. When institutional repository managers register their institutional repository with OAISTER, the main reason is that of making the institutional repository interoperable.

2.11 Proposals for dealing with securing contents for an institutional repository

Content recruitment for institutional repositories is by far the biggest challenge faced by institutional repositories. A strong recommendation by Pinfield, (2005) and Harnad (2006) is that an institutional *policy* should recognise that the deposit of materials is mandatory. According to Jantzi (2009), and Swan and Brown (2005), academic staff are of the view that should their employer include a clause in their job descriptions stating that research materials should be deposited in the institutional repository they would be more willing to do so.

2.11.1 Self-Archiving

According to SPARC (2011), self-archiving refers to

> "the practice of scholars depositing copies of their research papers in electronic repositories or open archives. It is a broad term often applied to the electronic posting, without publisher mediation, of author-supplied research. In practice, self-archiving encompasses both the posting of pre-prints and (in many instances) published papers by individual researchers on personal web sites and the inclusion of such research on the discipline-specific e-print servers... Repositories can either be institutionally-based capturing and preserving the intellectual output of a single or multi-university community or discipline-based as is the case of the most famous open repository, the physics arXiv" (SPARC, 2011).

Self-archiving has to be done by the authors themselves. This can only be effective and possible if the authors have been educated on the deposit process and also made aware of the benefits of open access repositories by institutional repository managers. It is of paramount importance for institutional repository mangers to encourage authors to self-archive. As mentioned above, a study by Harnad (2006) showed that 95% of researchers are of the opinion that they can only self-archive their work if their employer wanted it.

Self-archiving is beneficial to the scholars, researchers, institution and society as a whole. Some of the benefits of self-archiving that have been identified in the literature are:

a) "Papers that are freely available online are on average downloaded and cited earlier and more often than those that are not;
b) When researchers deposit their academic work in an open access repository, they increase their profile as author on a worldwide basis, thus increasing the impact of their research;
c) Regular submission of an author's work to a repository provides an author with a central archive of their work and a record of publications to add to their CV" (SPARC, 2011).

Although self-archiving is a good way of populating a repository, it has been found in the literature that significant barriers to self-archiving exist. The biggest barrier is the concern about copyright. Authors are concerned about receiving permission from their publishers because they do not want to infringe copyright and as a result they do not self-archive their work (Swan, 2006).

The second powerful barrier to self-archiving, as noted in the literature, is age. Younger faculty members have been found to self-archive a greater percentage of their research as compared to the older staff members (Kim, 2010). This is thought to be so because younger researchers are more familiar with disseminating their work on the internet than older ones. Younger faculty members are also found to be more computer literate and much more familiar with communication technologies.

The third barrier to self-archiving is the time and effort required to do so. Researchers have many time constraints and do not want to spend additional time and effort self-archiving (Kim, 2010).

2.11.2 Encourage deposit of staff publication

A strategy of encouraging staff to deposit their publications was found successful in helping to populate a repository. An example of such a strategy was found at the Harvard University in the Faculty of Arts and Sciences (Guterman, 2008).

2.11.3 Policies that mandate deposit

Another solution that would increase materials in a repository, is implementing a policy that mandates the deposit of student theses and dissertations. Postgraduate students from all the faculties of the institution would be required to submit electronic copies of their theses and

dissertations to the faculty administration offices before graduation. Many academic institutions make such policies for their Masters and Doctorate candidates (Nabe, 2010).

As mentioned above, another way to recruit content is to mandate staff to deposit their research articles. Cullen and Chawner (2009) argue that where institutions choose to have policies that mandate deposit, they should be flexible and not force staff to submit. They should try to always encourage staff to deposit whatever publication they wish to deposit. They further observed that in so doing, it encourages staff to deposit and as a result, the repositories in those institutions have high rates of deposit. Mandates help populate institutional repositories because the main problem institutional repositories are facing is content recruitment. According Cullen and Chawner (2009), many universities around the world have been successful in establishing institutional repositories but have failed to populate them. An example of a university, which has such a mandatory policy, is the Queensland University of Technology in Australia. The mandate policy was designed for academic staff to self-archive, and has been successful (Cullen and Chawner, 2009). Another example is the Harvard University in the USA (Cullen and Chawner, 2009).

2.11.4 Marketing the repository

Scholarly communication keeps changing from time to time and academics need to be educated on this aspect. This can be done in the form of promoting the new scholarly communication cultures to academics so that they can understand and adopt them in order to make them deposit their research publications in the repository. Promoting the repository can be dome in a number of ways, some of the marketing strategies from the literature include;

a) "knowing the audience;

b) content of the marketing message; and

c) delivering the message" (Nabe, 2010: 62).

For each topic, there are questions that the institutional repository team should focus on. These questions will also guide them to formulate a powerful message to their audience. According to Nabe (2010) a "marketing plan includes two aspects which are crafting a message and finding

the best way to deliver it" (2010: 63). This message must be appealing and also address the needs of the targeted community. The above three marketing strategies are discussed in depth below.

2.11.4.1 Knowing the Audience

The first strategy requires the institutional repository team to do a survey. The aim of the survey is to gather the information needs of the institutional community and how they perceive the repository.

2.11.4.2 Content of the marketing message

The second strategy involves the elements of a marketing message as stated by Nabe (2010). This will include the following aspects:

- The definition of institutional repository
- The reasons the institution is running one
- Listing both the benefits of an institutional repository and the reasons the individuals should contribute. This part is the core of the marketing plan where the contributors get to understand why they should participate in the institutional repository. The team that does the marketing should teach and educate the contributors the benefits of an institutional repository which includes; increased access; more visibility and better preservation.
- Those stakeholders and individuals who will be interested in participating in the repository will contact the marketing team for any enquires and clarification on what to do.

2.11.4.3 Delivering the message

The third and last strategy is when the institutional repository team delivers the message to the rest of the institution's population. They could target the institution's publications such as popular campus newspapers, newsletters, research newsletter and announcements in the institutions radio or television station. In addition, the institutional repository could make their own promotional materials. This can include brochures, fliers, and posters around campus, which clearly show the name of the repository and relevant contact information, i.e. telephone numbers, an email address and website URL link (Nabe, 2010). Branding the institutional repository is

very important. A logo and a name will distinctively distinguish it from other clubs and organisations found in the institution.

2.11.5 Barriers of contents recruitment for an institutional repository

One of the barriers cited in the literature is that many scholars are unaware of the open access movement or the existence of the institutional repositories. The second barrier is that many scholars do not understand the benefits provided by open access or the institutional repositories. The third barrier is that researchers are not aware of the potential for institutional repositories to serve their research interests and needs. Furthermore, outreach efforts by institutional repository implementers do not speak their language or address their problems. They prefer spending time doing activities that will bring them professional recognition and add value to their research, i.e. authorship and collaborations. The fourth barrier is that some researchers worry that open access will lead to plagiarism (Troll Covey, 2011).

In summary, one could argue that institutional repository implementers should take into consideration technical and logistical assistance for self-archiving. This can be done by hiring clerical assistance to deposit materials on behalf of researchers who lack the necessary computer skill, allowing them to participate (Kim, 2010). This support may also alleviate researchers' concerns about the extra time and effort inherent in self-archiving, which are barriers to self-archiving. Institutional repository implementers should clearly address the concern and confusion over copyright issues. This can be done by providing services for copyright management that can assist researchers in determining the legal implications of self-archiving their publications. Dealing with these issues has financial implications, which may also be regarded as a barrier.

2.12 Software used for institutional repositories

There are a number of open source and proprietary software packages available for institutional repository development. Below are some of the better-known software packages offering institutional repository software, listed in alphabetical order in the table below:

Open Source Software	*Proprietary Software*
1. Archimede	1. Content Pro –Innovative
2. CDSware	2. CONTENTdm
3. DSpace	3. Digital Commons, formerly Bepress
4. EPrints	4. Digitool
5. ETD-db	5. Documentum
6. Fedora	6. Open Repository
7. Greenstone	
8. I-Tor	

Table 1: Open source and proprietary software

According to the Directory of Open Access Repositories (DOAR), the two most popular open source repository software systems in use are DSpace and ePrints. The third open source repository software system in use is Fedora as shown by figures from the Fedora Commons (OpenDOAR, 2009, Smith, 2003). Eprints software as indicated by Ball (2010) is the most broadly distributed, and installed base of any of the repository software systems. It was developed at the University of Southampton and it was released in 2001. DSpace, which is discussed below is the second highly used software.

DSpace is easier to customise than the other two and many academic institutions use it. According to Ball (2010)

> "DSpace design supports the participation of schools, departments, research centres, and other units typical of a large research institution. As the requirements of these communities might vary, DSpace allows the workflow and other policy-related aspects of the system to be customized to serve the content, authorization, and intellectual property issues of each" (Ball, 2010:11).

Supporting this type of distributed content administration, coupled with integrated tools to support digital preservation planning, makes DSpace well suited to the realities of managing a repository in a large institutional setting. One reason why institutions tend to prefer it is that DSpace is focused on the problem of long-term preservation of deposited research material. This

allows DSpace users to offer services both for hosting institutional repository content and maintaining the content for archival periods (Ball, 2010).

The third most popularly used open access institutional repository is the Flexible Extensible Digital Object and Repository Architecture (FEDORA). Fedora software was developed by the University of Virginia in cooperation with Cornell University Information Science. It was funded by the Defense Advanced Research Projects Agency (DARPA) and the National Science Foundation (Markey et al., 2007). The project also received grants from the Andrew Melon Foundation. Fedora software can be used for institutional repositories, library collections management, multimedia authoring systems, archival repositories, and digital repositories. When Fedora's open-source software is downloaded for use by institutions, the institution needs to have staff with technical expertise to customize it (Markey et al., 2007).

One important consideration to be observed in institutional repository establishment is that as open access repository software can be obtained and installed for free, and can run on a basic hardware configuration, it is important to have technical staff to assist in installing the repository software. Funds may also need to be expended on improving hardware, as "disk storage, server capacity, and perhaps other specifications would need to be upgraded as the repository moves from a pilot stage into public operation and heavy use" (Ware, 2004:116). Most labour costs, though, will relate to non-technical staff. In particular, staff will need to be trained in the on-going management of the repository, which includes clerical staff who would be assisting authors to deposit their work and checking copyright permissions from publishers. Money may also be spent on advocacy and marketing the repository (Ware, 2004). There is a general agreement that there is now adequate, easily available, software to create and maintain an institutional repository. As a result, the challenges in setting up an institutional repository are now seen as being less to do with technology and are more to do with managerial, organizational and cultural issues (Ware, 2004).

ePrint software is the dominant software in Europe as it was developed in the UK (Barley, 2006). In the USA the dominant software used for institutional repositories is DSpace, developed by Hewlett Packard and MIT Alliance and Fedora, developed by Virginia Tech (Markey et al., 2007). ARROW is most widely used in Australia, and Archimede and DSpace are considered the

main software in Canada (Shearer, 2006). The challenge for new institutions in Africa is not to develop their own, but to identify with the dominant software that matches their interests.

2.13 International developments on institutional repositories

The developments of institutional repositories have largely taken place in universities in Europe, America, Australia, Canada and more recently in Africa, especially in South Africa. Below is an analysis and discussion of international developments on institutional repositories.

2.13.1 Institutional repositories in the UK

The movement for open access and the establishment of institutional repositories developed in the United Kingdom (Drake, 2004). The European Union countries formed a Consortium of University Research Libraries (CURL) and the Joint Information Systems Committee (JISC). CURL's mission was to increase the ability of research universities to share research for the benefit of research communities, while JISC aimed to support teaching, learning, research, and administration in higher education through the use of information and communications technology.

Ware (2004) reviewed developments of institutional repositories, and indicated that JISC funded a series of projects under the Focus on Access to Institutional Resources (FAIR). One of these projects was Securing a Hybrid Environment for Research Preservation and Access (SHERPA) project, which was set up mainly to inspire change in the way universities disseminated their scholarly research outputs. This change was about encouraging institutions to establish open access institutional repositories using the ePrint software, which became the dominant software in Europe. The SHERPA project was also aimed at informing JISC on issues such as how institutional repositories can be built and maintained, how they tackle intellectual property rights (IPR) and copyright, and how to market institutional repositories.

The SHERPA project proposal was produced under the support of CURL. It provided substantial financial support and covered key elements of the digital preservation costs and some hardware costs (Johnson, 2007a). The SHERPA Project aimed to establish institutional repositories in leading UK universities. This would comply with the Open Archives Initiative Protocol for

Metadata Harvesting (OAI-PMH) using ePrints software. It also aimed to carry out detailed investigations into digital preservation in institutional repositories. The institutional repository projects that JISC and CURL carried through the SHERPA Project, supported the goals of both organizations and promoted collaborative development and operations (Ayris, MacColl and Pinfield, 2002). Below are examples of institutional repository projects carried out in the UK. CURL's initiative played an important role in helping universities in the UK to think about the benefits of establishing institutional repositories.

The executive committee of SHERPA was based at the University of Nottingham. The committee worked with the partners to carry out the projects by using the ePrint software. In addition to the above-mentioned partners, seven institutions provided assistance demonstrations on workflows, processes, staffing models, approaches and implementations. The SHERPA project benefited from the outset from the involvement of two different models of repository organisations. This was at individual institutions and a consortium approach adopted by the White Rose Group of Leeds, Sheffield, and York (Johnson, 2007b).

The second FAIR programme project was based in Bath University and focused on developing a national service provider repository of ePrint records (Ware, 2004). The third and fourth projects were based at the University of Glasgow respectively. The DAEDALUS project mainly focused on the establishment of OAI-PMH compliant digital collections. The fifth project dealt specifically with copyright and was called the Rights Metadata for Open Archiving (RoMEO) and was a JISC-funded project (Project RoMEO, 2003) at Loughborough University. It was to investigate the rights issues surrounding the self-archiving of research in the UK community under the Open Archive Initiative Protocol for Metadata Harvesting. By surveying the academic community, it ascertained how research literature and metadata was used and how it should be protected. From this work, the RoMEO project created a list of publishers' conditions for self-archiving to eliminate problems (RoMEO, 2011).

The Digital Repository Infrastructure Vision for European Research (DRIVER), was a SHERPA project that was designed to help the institutions and universities in the UK and around the European countries to establish their institutional repositories. This project was also a way to check if the open access message would be successful in the European Union region and encouraging them to do research (Johnson, 2008).

A complementary project to the RoMEO service, provided by SHERPA for authors and repository administrators, was carried out and was named the JULIET project. It detailed the requirements from research funders' mandates as they pertained to research outputs, allowing publisher rules and funder requirements to be viewed alongside each other and for the information to be linked to the RoMEO site (Johnson, 2007c).

After all the projects were complete, SHERPA started the advocacy project. Advocacy deals with building an informed awareness for which a core message and ethos is essential. This is because many academic libraries in the UK had implemented institutional repositories and SHERPA decided that it was time to advocate for their use. For an institutional repository to be successful, cultural change must be achieved through advocacy, which can be achieved through "getting the right message to the right people with the tone and content varied by audience" (Johnson, 2007c: 23). Although an institutional repository may have funding, staffing, fundamental values and policies, it still needs advocacy.

The key questions, which libraries with institutional repositories must try to answer, are "what the repository does and who is it for?" (Johnson, 2007c: 23). SHERPA observed that neglecting advocacy results in repository decline and stagnation, but with the right level of engagement across the institution, advocacy becomes the key to success for the SHERPA Partners' IRs. It produced "seven (and a half) pillars of advocacy" (Johnson, 2007c: 23) which are discussed in detail in chapter five (cf. 5.3.4.2). Advocacy has been the heart of the SHERPA Partner activity and the cornerstone of achieving cultural change within institutions (Markland and Brophy, 2005).

For a repository to be established, staffing, software, maintenance, content recruitment and equipment are needed (Gibbons, 2004). Open Access is a truly global movement with benefits for researchers and institutions. In the United Kingdom SHERPA has applied experience across its partnerships to develop useful guidance and support (Johnson, 2007a). The SHERPA project has paved the way for institutional repositories to be seen as an important part of research and archiving of academic research materials in Europe.

2.13.1.1 Institutional repository challenges faced in the UK

Although the UK is a leader in the movement towards open access publishing and the implementation of institutional repositories, a number of obstacles were encountered. The major obstacle that still exists is the recruitment of content into institutional repository. Researchers have found that institutional repositories are faced with difficulties in being populated. Swan & Awre (2006) reported that even though almost one third of UK universities have open access institutional repositories, they do not have much content, especially research such as working papers, pre-prints and theses. Swan & Awre (2006) attribute this low content in institutional repository to a lack of knowledge or information of open access by researchers. It seems that researchers are not aware of Open Access at all, or are aware but not informed of its benefits, or are aware of its benefits but not willing to participate. This problem is partly due to the authors' lack of understanding when it comes to obtaining copyright permissions from publishers to deposit. Researchers have questions about who will use the materials that they deposited, how it will be used, and whether it will be appropriately attributed. In order to become embedded in the culture of research-based organisation, both the habits and procedures of the researchers, academic policies and institutions involved will have to change (Swan & Awre, 2006).

2.13.2 Institutional Repositories in the USA

In the USA, a number of organisations such as the Association of Research Libraries (ARL), the Association of College and Research Libraries (ACRL), the Coalition of Networked Information (CNI), the Association of American Universities (AAU) and SPARC America support institutional repositories as part of their efforts to reform scholarly communication and achieve open access to publicly funded research. ARL has worked since 1970 to gather and disseminate the best practices for library needs (Bailey, 2006). As part of its commitment, the ARL maintains an active publications program best known for its SPEC Kits. Through the Collaborative Research and Writing Program, librarians work with ARL staff to design SPEC surveys and write publications known as SPEC Kits. These kits were originally established as an information source for ARL member libraries. Since then, the SPEC series has grown to serve the needs of the library community worldwide. The Coalition for Networked Information (CNI) is an organization whose mission is to promote networked information technology as a way to further

the advancement of intellectual collaboration and productivity. It is a joint initiative of the Association of Research Libraries (ARL) as well as EDUCAUSE (Bailey et al., 2006).

The movement towards the open access model of publishing and establishment and development of institutional repositories gained prominence in the United States after two big institutional repositories were launched in 2002 at the University of California and at the Massachusetts Institute of Technology (MIT). The leadership of MIT, in collaboration with the Hewlett Packard Corporation, developed and introduced DSpace software in 2002. The Andrew W. Mellon Foundation funded MIT's DSpace to be duplicated in several other universities around the world, as it is open source software that other universities could use to develop and implement their own repositories (Lynch, 2003). Since then an increasing number of research libraries and their parent institutions have established institutional repositories to collect and provide access to diverse digital materials. Afterwards a number of projects attempted to find out how institutions in America were dealing with the new developments of scholarly publishing.

In 2006, Bailey (2006) conducted a survey, which showed that many American universities, especially since 2004, were establishing institutional repositories. He found that the top three reasons given by the implementers was that of increasing global visibility, as well as preserving and providing free access to the institution's scholarship (Bailey, 2006).

Several studies were carried out between the years 2005- 2007 under the MIRACLE (Making Institutional Repositories a Collaborative Learning Environment) Project, which were aimed at investigating the implementation of institutional repositories at academic institutions in the USA. This project was meant to identify models and best practices for the administration, technical infrastructure, and access to digital collections. The main objective of the project as stated by Markey and others (2007) was to identify specific factors contributing to the success of institutional repositories and effective ways of accessing and using institutional repositories. For example, one of the MIRACLE projects involved a comparison across institutions at different stages of institutional repository development in USA, which tried to include all research institutions and universities.

While the developments in America by MIRACLE are similar to what the SHERPA and FAIR did in the UK, it appears there are differences, namely the presence of more organisations that

provided funding in America than those in the UK. Companies such as Hewlett Packard helped in steering or influencing these software developments. The obvious deduction from the experience in America was that there was a large flow towards institutions establishing institutional repositories through the influence of many organisations advocating their establishment and providing funding.

2.13.2.1 Institutional Repository challenges faced in the USA

The challenges that were faced by the American universities after the establishment of institutional repositories are similar to those experienced by institutions in the UK, and which researchers all over the world have identified. These challenges included content recruitment, staffing and faculty awareness. As in the UK, the challenge of copyright issues was also noted in the USA.

This was confirmed by Bailey who noted that "among the top three challenges that respondents face in implementing, promoting, and running an institutional repositories are content recruitment/building a critical mass of content, staffing and faculty awareness and engagement" (2006: 20). In America, challenges included resources and integrating the staff and workflow of institutional repositories with structures (Bailey, 2006).

2.13.3 Institutional Repositories in Canada

Institutional repository projects started in Canada in 2003. They were initiated by the Canadian Association of Research Libraries (CARL) who also launched the CARL Institutional Repository Project. This project was meant to provide support for Canadian academic or research institutions as implementers of institutional repositories. The CARL working group developed a vision for institutional repositories in Canada. This included setting priorities and specific projects for the CARL Institutional Repository Program and coordinating the implementation of projects. Mark and Shearer (2006), the head of CARL institutional repository project, presented the following rationale for the project at an Open Access conference hosted by the CARL in 2003:

> "It is the responsibility of each academic institution to preserve, organize and distribute the intellectual output of their faculty". Furthermore, institutional repositories are "a way that

institutions can 'get back' some of the output from the researchers and accelerate the movement toward open sharing of knowledge." (Mark & Shearer, 2006:3)

This marked the beginning of institutional repository establishments in academic and research institutions in Canada. Another project in Canada, which was an institutional repository initiative, was when CARL worked with the Canadian Institutes of Health Research (CIHR) to develop a coordinated communication strategy for the CIHR Policy on Access to Research Outputs. The policy was published in 2006 and was commended by SPARC. The policy was to be used to govern peer-reviewed journal publications, research materials, and final research data stemming from CIHR funding (McLennan and Mark, 2006). It also marked a significant step forward for Canadian science and it put Canada in the forefront of the global open access movement. According to McLennan and Mark (2006), CIHR is the major federal agency responsible for funding health research in Canada.

Another initiative in Canada was the interoperability project between institutional and disciplinary repositories and CARL. The two were conducting a feasibility study on the implementation of the Simple Web-service Offering Repository Deposit (SWORD) protocol to enable simultaneous deposit into multiple repositories. The study was meant to survey CARL libraries to determine interest in implementing SWORD. This was to be followed by a Pilot Project to implement the protocol at interested repositories (Markey and others, 2007).

2.13.3.1 Institutional Repository challenges faced in Canada

Developments in Canada are similar to those in USA, Europe and Australia, as the phenomenon of institutional repository is worldwide. In Canada even with a variety of creative ideas and promotional activities, faculties were reluctant to participate. Mark & Shearer (2006) state that given the vision and the potential of institutional repositories, it is surprising and disappointing that institutional repository collections have generally grown more slowly than proponents had anticipated. Just as in America and Europe, Canada implementers have found that recruiting content is the biggest challenge and frustration.

In the literature as cited by (Swan, 2004; Swan and Brown, 2005; Mark and Shearer, 2006) a number of reasons are presented why faculty participation rates are so low. At the most basic level, faculty members lack awareness of the existence of institutional repositories. Several

surveys have found that many academic authors are not familiar with the concept of any institutional repositories on campus (Swan, 2004; Swan and Brown, 2005; Mark and Shearer, 2006).

Canada had a problem of metadata inconsistency and incompleteness. This challenged the effective harvesting and searching of institutional repository records for CARL members, which was not found in other countries. The problem was attributed to the wide variety of content types that were used. This often resulted in variable uses of unqualified Dublin Core elements. The CARL project members decided to solve it by better harmonizing the metadata of participating institutional repositories. Members of the Canadian institutional repository community raised the issue at a meeting held during at a Workshop in Halifax in October 2003. In response, CARL began to develop a National Metadata Profile for institutional repositories in the CARL website (Mark & Shearer, 2006).

2.13.4 Institutional Repositories in Australia

The concept of developing and implementing institutional repositories in Australia started because many universities and research institutions were becoming interested and aware of the importance of collecting, preserving, reusing digital content produced for research, teaching and learning. They then became interested in establishing institutional repositories to house the digital information which their staff and students were producing, as they discovered that these research outputs were institutional assets. The interest was also sparked by the new information technologies that offered opportunities for collaboration, flexible and online learning. The Australian academic libraries responded to the challenges of new technologies by taking the opportunity of redefining their fundamental role in the creation, distribution and provision of access to information. This also followed the international open access and institutional repository movement, in both the UK and the USA as discussed above.

In June of 2003, the Australian Commonwealth Department of Education, Science and Training (DEST) called for proposals from different universities and institutions in Australia with regard to research on how it is discovered, created, managed and disseminated in a digital environment (Treloar, 2004). The call for proposals was meant to do three things. First, it was to bring all stakeholders together in order to fund and help promote research. Secondly it was meant to build

the information infrastructure through the development of shared digital repositories and technical services. Lastly, it was meant to oversee the new developments of establishing institutional repositories. Out of the 14 projects that were submitted after the proposal call four projects were funded. These are:

- "The Australian Research Repositories Online to the World (ARROW),
- Australian Digital Theses Program Expansion and Redevelopment (ADT),
- Towards an Australian Partnership for Sustainable Repositories (APSR) and
- Meta Access Management System (MAMS)" (Treloar 2004).

The projects were named the *Federated Repositories of Digital Objects* (FRODO). They were funded for a period of three years and were to commence at the start of 2004 (Treloar, 2004). Many top universities throughout Australia had already established digital repositories that were linked to form ARROW (Australian Research Repositories Online to the World). ARROW was developed and funded by the government to establish a number of institutional repositories. It also tested the searching of aggregated metadata harvested from the developed repositories (Treloar, 2004). The ARROW Discovery Service, developed and operated by the National Library of Australia, provided the search interface.

The ARROW project was a collaboration of three universities, Monash University (lead institution), the National Library of Australia, the University of New South Wales and Swinburne University of Technology. This project was approved in 2003, with the funding covering three years until December 31, 2006. The main task of the ARROW project was to provide software applications for institutional repositories (Treloar, 2004).

2.13.4.1 Institutional Repository challenges faced in Australia

From the literature, it is clear that the worldwide challenges for implementing and establishing institutional repositories are similar. In Australia, there was a challenge of institutional repositories being seriously underutilised as observed by Whitehead (2005).

2.13.5 Institutional Repositories in Africa

African universities, as in other parts of the world, are beginning to embrace the initiative towards establishing open access institutional repositories. This was illustrated by the number of African countries that have established repositories in their institutions, despite their challenges. Although "developing countries, such as South Africa, tend to think that they themselves have very little to offer in terms of knowledge creation and transfer" (Van Deventer and Pienaar, 2008:1), they assert that this is not the case. It is by creating and maintaining repositories that the opportunity to learn, teach and to participate in eResearch is provided in Africa.

In December 2008, there were twenty institutional repositories in Africa (Christian, 2008). This was a result of awareness campaigns through a series of workshops on the use of open access institutional repositories. According to OpenDOAR in February 2010, there were 35 institutional repositories in Africa out of which 24 were found in South Africa, two each in Egypt and Kenya and one each in Botswana, Cape Verde, Ethiopia, Ghana, Namibia, Nigeria, Uganda and Zimbabwe (Asamoah-Hassan, 2010). In West Africa, KNUSTSpace in Ghana is a pace setter in the sub-region, being one of the three and the one with the highest number of entries in February 2010 (Asamoah-Hassan, 2010). In Southern Africa, the University of Pretoria is the leading example with two repositories, one for electronic theses and dissertations, and the other for their research materials (Hammes, 2010). This shows that that there is a growing understanding of the usefulness of institutional repositories in Africa and also the acceptance of institutional repositories that make local research available globally.

A solution to the visibility of African research output is therefore through the development and establishment of open access institutional repositories by institutions to make their research output widely available and accessible to the outside world via the World Wide Web. In September 2011, the 14th Symposium on Electronic Theses and Dissertations was held for the first time in South Africa, Cape Town and more African universities were in attendance. With this kind of conference, it is hoped that more African universities will establish repositories and share their research output with the world to make African research visible.

2.13.5.1 Institutional Repository challenges faced in Africa

In Africa, the existing institutional repositories were the result of institutional initiative rather than driven by representative organisations. In some cases in Africa, the regional blocks started lobbying for the formation of committees that could spearhead the institutional repositories. For example, in South Africa, the South African Research Information Services (SARIS) was formed in conjunction with the Academy of Science of South Africa (ASSAf) to lead the open access movement which by encouraging institutions to establish institutional repositories (Van Deventer and Pienaar, 2008). The Electronic Information for Libraries (eIFL) and the Mellon Foundation, which had been playing an important role in the development of the South African information industry, played an important role in helping South African universities establish their repositories.

Christian (2008) outlined a number of challenges in Nigeria in the establishment, development and implementation of institutional repositories. These are similar to those found in Ghana by Asamoah-Hassan (2007), and noted elsewhere in Africa by Anbu (2006) and Moahi (2009). It is evident that Africa has some of the biggest problems and challenges. Some of the issues that affect the development of institutional repository in Africa include:

- The lack of awareness of the advantages of open access publishing in academic and research institutions in developing countries,
- inadequate information and communication infrastructure,
- lack of funding,
- poor advocacy for open access and
- copyright management issues (Christian, 2008: 28)

For any institution to be able to establish an institutional repository, it is very important to have funding. It is of paramount importance in acquiring hardware and human resources, since software is open source. In African countries, some academic and research institutions face serious problems with ICT infrastructure, power and expensive internet bandwidth which they cannot afford. Compared to other parts of the world, Africa does not have many funding

organisations. Electronic Information for Libraries (eIFL) helps African institutions from the continent but it is still work in progress. Although African countries face many challenges, one can argue that the establishment of institutional repositories will ensure that their national research will become part of the mainstream and contribute on an equal footing to the global knowledge pool.

This discussion has shown that acquisition of content in institutional repositories is still the central issue in all continents. The study has also observed the importance of organisations such as CURL in the UK, ACRL in USA, CARL in Canada and ARROW in Australia. These were actively involved in funding and helping universities establish institutional repositories, which does not seem to be the case in Africa.

2.13.5.2 The status of institutional repositories in South Africa

According to Van Deventer and Pienaar (2008), a national research and development strategy for South Africa was published in 2002. The strategy invited all stakeholders in the national innovation system to reconsider the renewal of the information services sector. This brought about the SARIS Project, which would investigate ways in which South African research institutes and university libraries could access expensive world research literature. This was a sign that eResearch was emerging in South Africa. It was recommended that a framework for eResearch services to the entire South African research community should be created, so that eResearch development and innovation services could be jointly funded as projects. Those who made use of the service would fund identified projects that proved to be essential to the research community. The whole system was to be co-ordinated at country level (Van Deventer and Pienaar, 2008: 2). Unfortunately, this did not materialise due to lack of national coordination. This meant that individual institutions had to start their own initiatives. Institutions with the help of eILF started establishing institutional repositories.

Currently, the OpenDOAR website shows that the number of institutional repositories in South Africa has grown from 14 in 2008 to 24 in 2011. Below is a list of South African repositories, both from the academic institutions and research organizations, which have been downloaded from ROAR and OpenDoar 2011.

2.13.5.3 Repositories in South Africa: 2011

From the list below, UP will be used for a comparative analysis with UCT in the discussion found in chapter 5. According to UP's Electronic Theses and Dissertations Repository website, it has 6621 theses and dissertations and has improved its rating in the international Webometrics ranking. In 2010, it was at number 65 worldwide in terms of content, size and visibility, and it was celebrating 10 years of existence.

Name of repository	Number of records in 2011	Software used
1. African Higher Education Research	165	AHERO
2. Boloka Research Repository: North West University	4002	DSpace
3. CSIR Research Space	4250	DSpace
4. Digital Innovations SA	Not stated	DSpace
5. Digital Knowledge	1598	Digital Commons
6. Durban University of Technology IR	248	DSpace
7. Rhodes eResearch Repository	1279	ePrints
8. Scientific Electronic Library Online- SA	22	SciELO
9. Stellenbosch University- SUNScholar Repository	3390	DSpace
10. UCT Computer Science Research Archive	290	ePrints
11. University of Cape Town Law Space	222	DSpace
12. University of Fort Hare Institutional Repository	121	DSpace
13. University of Free State ETD	Not stated	ETD-db
14. University of Johannesburg DigiSPace	405	DSpace
15. University of Kwa-Zulu Natal Research Space	3882	DSpace
16. University of Limpopo	70	DSpace
17. University of Pretoria – UPSpace + Open UP	13864	DSpace
18. University of Pretoria Electronic theses and Dissertations	6621	ETD-db
19. University of South Africa Institutional Repository (Unisa)	4097	DSpace
20. University of the Western Cape Institutional Repository	26	DSpace
21. University of the Western Cape Theses and Dissertations	1525	Chisimba
22. University of Witwatersrand (WIReDSpace)	3787	DSpace
23. University of Witwatersrand Institutional Repository	5405	DSpace
24. University of Zululand	765	DSpace

Adapted from ***Directory of Open Access Repositories 2011:*** http://www.opendoar.org/countrylist.php#South Africa

(The records above were updated on the 28th November 2011)

2.14 Metadata assignment in a digital institutional repository

Metadata is the information that describes the material deposited in the repository and is defined as "data used to describe other data" (Caplan, 2003: 1). Metadata explains, locates, and describes

an information resource. It also helps in the easy retrieval, management and use of these resources. Metadata of records accompany all the digital objects stored in an institutional repository. In the library environment traditional library cataloguing, MARC 21, AACR2 are all metadata standards. For example, metadata includes such details as the title of the article deposited and the name of the author. Usually, metadata is entered into the repository at the time of deposit. It is important to have standardised metadata formats as they enable users to search easily and effectively and to find and retrieve information from the repository (Pappalardo and Fitzgerald, 2007). For example, an end user should be able to conduct a search for '*Khumalo, Betrice'* and retrieve many articles written by authors with the name '*Khumalo.*' The metadata describing the content of the repository and not the actual content itself is searched, because searching the actual content would take too long. According to NISO (2004: 2) "metadata is key to ensuring that resources will survive and continue to be accessible into the future". This means that metadata is very important for resources discovery. Other importance's of metadata is that it helps in organising electronic resources, gives resources digital identification, helps in interoperability, it also supports archiving and preservation (NISO, 2004).

2.14.1 The importance of metadata in a repository

The most important use of metadata in an institutional repository is to help users to retrieve resources, and to facilitate discovery of repository content. Below is a list of the functions of metadata in an institutional repository:

- "helps users identify resources,
- helps repository managers organise content,
- brings similar resources together,
- distinguishes similar resources,
- gives location information,
- facilitates harvesting of repository content by external systems or search engines, and
- supports archiving and preservation" (McGill, 2011: 73).

2.14.2 Metadata creation

Metadata is created by a number of people who are involved in the repository. These include the authors who create metadata for their research articles when submitting them. It also includes the institutional repository manager or cataloguer when editing or adding the item into the repository. Authors also create metadata of the institutional repository, and metadata can be taken from contents that have been deposited in the repository by use of automated tools. Cataloguers and institutional repository managers mainly add administrative and bibliographic information to resources such as subject headings (McGill, 2011).

2.14.3 Metadata use

Metadata has to be stored as it manages a resource. It can also be used to check the currency of file formats and to review the strength of the repository (McGill, 2011). In order for external services, which are mainly search engines, to access metadata records for repositories and enable retrieval of resources, standards such as Open Archive Initiative, the Protocol for Metadata Harvesting (OAI-PMH) should be used. These standards create interoperability connecting distant institutional repository content through search engine capabilities.

OAI-PMH was designed to harvest metadata and associated resources that are distributed across different OAI-compliant servers. It is a key building block that makes the institutional repository the platform of choice to support open access to institutional research output and, to connect all distributed servers into a seamless global digital library (Pappalardo and Fitzgerald, 2007). The accessibility and impact of materials is greatly reduced if institutional repositories remain invisible to others because of the lack of interoperable standards. The OAI-PMH is now widely adopted by libraries, publishing, and scientific communities in order to ensure that their resources on the Web are interoperable with each other. It is based on the Dublin Core metadata standards. Most open source digital repository software such as DSpace, Fedora, and ePrints all support the OAI-PMH, which in turn allows the digital repositories to be harvested by search engines like OAIster and Google Scholar (Amaral, 2007).

In summary, interoperability is created by OAI-PMH standards when distant institutional repository content is connected through search engine capabilities. The ability to search content

is limited when institutional repositories do not adhere to these standards. These can make institutional repositories only accessible to community members who are locally aware of specific content. Westell (2006) states that interoperability indicates an ability and openness on behalf of the institution or the library that is willing to contribute to national and international scholarship.

2.14.4 Metadata Standards

There are many different metadata standards that have been developed by the archives community. These include the Dublin Core metadata element set, Metadata Object Description Schema (MODS), Visual Resources Association core (VRA), the Encoded Archival Description (EAD), and the Metadata Encoding & Transmission Standard (METS). The most commonly used standard, and the one required for interoperability and harvesting via the OAI-PMH in institutional repositories, is the Dublin Core. It was developed at a workshop held in Dublin, Ohio in 1995 by the Dublin Core Metadata Initiative (DCMI) (Caplan, 2003).

The Dublin Core standard has 15 elements or fields, which are optional and repeatable, and may appear in any order. These fields are:

- "Title of the article/material;
- Creator – the name(s) of the author/s;
- Subject matter of the material;
- Description of the content of the material;
- Publisher;
- Contributor – a person or corporate contributor to the material's content;
- Date that the material was created, modified or made available;
- Type – the nature or genre of the material's content;
- Format in which the material appears;
- Identifier – an unambiguous reference to the material within a given context. Usually material is identified by a string or number conforming to a formal identification system. For example, the International Standard Book Number (ISBN). Metadata is not limited to these baseline requirements;
- Source;

- Language;
- Relation;
- Coverage and
- Rights" (McGill, 2011: 75).

Other fields of metadata can be imposed by the repository. These fields may indicate, for example, whether the material has been peer reviewed, whether the full text of the material is available, and the status of the material, i.e. published, in press, or unpublished (Pappalardo and Fitzgerald, 2007). Dublin Core is compliant with OAI - PMH harvesting and therefore, all institutional repository solutions which use Dublin Core as their metadata standard are compatible with OAI. Metadata is also important in supporting repository management such as administration and preservation (Aschenbrenner & Kaiser, 2005).

2.14.5 Types of metadata

There are three main types of metadata that can be used in repositories, which are discussed below.

Descriptive metadata: This type of metadata is used for a number of purposes such as identification, selection, collocation and acquisition (Caplan, 2003). A common example of descriptive metadata is the library cataloguing and the Dublin Core schema. Descriptive metadata will include resource title, author, and keywords.

Structural metadata: According to Caplan (2003: 5), structural metadata can be thought of "as glue that holds compound digital objects together." This means that it is used to describe the relationship between items. For example, a book has many chapters, which are made up of pages. Structural metadata will show the relationship between the physical files and pages. It will also show the relationship between the pages and the chapters and lastly, the chapters and the whole book.

Administrative metadata: This is mainly used by information managers to help manage the resource. It includes the information about when and how the resource was created, who is responsible for controlling access and what restrictions are there for the resource (Caplan, 2003).

2.14.6 Important points to note about metadata

Pappalardo and Fitzgerald (2007) noted that one other important factor to take into consideration is that when entering metadata into a repository, it must be done accurately and correctly, or users conducting a search might not find the resource that the metadata is meant to describe. Academic authors often seem reluctant to provide high-quality metadata, because the process of entering metadata is tedious and time consuming (McGill, 2011). It is advisable that the repository manager, or another staff member, is responsible for checking the quality of metadata. This ensures that it is sufficiently detailed and accurate. The administrator must approve an item before it can become accessible. The administrator should accept, edit or reject a submission and should also make contributing authors aware that it is essential for them to enter accurate metadata so that their work can be easily found and accessed by end-users. When installing the software, the technician would need to configure the metadata formats and customise the user interface. When this is done at the outset it is easier to do. It is more complex to change it once the uploading of materials has started. Those in charge of metadata should ensure that they are satisfied with the metadata formats selected before uploading too many documents.

Chan (2004) has observed that "the highly distributed nature of resources scattered across the World Wide Web has made meaningful searches difficult and often elusive" (2004: 281). The difficulty in finding what one is looking for stems from the reason that it is often very difficult to establish the origins and authenticity of Web resources from the general search engines. The reason is that digital resources often lack a proper description of what the resource is about due to a lack of metadata (Chan, 2004).

2.15 Institutional repository and the Copyright issues

Copyright can be defined as "a collection of legal rights that are attached to an original work when it is created" (Pappalardo and Fitzgerald, 2007:69). It allows the copyright owner to

prevent others from, for example, copying or downloading their material without permission. In so doing the copyright holder would be protecting his/her material.

When considering copyright, the author of any piece of work is the first owner of copyright. This is the basic principle of copyright ownership. Secondly, copyright ownership can be assigned to other people such as publishers. Many publishing houses normally require authors to assign copyright to them before they can publish their work. The third point to note is that employers or institutions also take ownership of their workers' or students' copyrights (Pappalardo and Fitzgerald, 2007). For example, at UP, the rights with regard to intellectual property that is produced by a student during his/her studies vest in the University, including the copyright to the UP's theses and dissertations (Olivier & Louw, 2011). This also applies at UCT and many other universities in the world.

An institution's open access policy should stipulate that a depositor must own copyright of the work being submitted to the repository, or must have permission from the copyright owner to deposit it (Pappalardo and Fitzgerald, 2007). This differs from one institution to the other, depending on their open access policy. In situations where the institution is the copyright owner of the material, it will be easier to mandate the deposit of material in the institutional repository, as there will be no issues with regard to obtaining permission from publishers or other copyright owners.

Issues about copyright will need to be explained in full and clarified to authors by the institutional repository managers in order for authors to know if they can self-archive their work in an institutional repository or not. Institutional repository managers will also have to request authors to check their publisher's website for terms and conditions about submitting work in an institutional repository before they self-archive, so that agreements are not breached. Authors who are not copyright owners may need to seek permission from their publisher before depositing their materials in an institutional repository.

2.16 Challenges in establishing institutional repositories in Africa

A number of challenges have been highlighted in the literature by several authors, namely Swan and Brown, 2005; Thomas, 2007; Moahi, 2009; Kingsley, 2008; and Lamptey and Corletey, 2011, with regard to establishing institutional repositories. Among these are inadequate funding, inadequate power, ICT connectivity and infrastructure, technological factors, legal or copyright issues, lack of advocacy, lack of faculty involvement, and technical issues. Below is a discussion of each challenge.

2.16.1 Inadequate funding

Many authors and researchers who have written about the establishment of institutional repositories have all noted that, in Africa the biggest challenge to establishing institutional repositories is that of inadequate funding by universities or research organisations. Although software is free, the cost of acquiring hardware (such as scanners and servers), subscription fees for internet bandwidth, the cost for maintaining the hardware and human resources are all very high. Institutions in Africa are also struggling to deal with budgets cuts from their governments (Christian, 2008). Therefore, obtaining funding for starting a project to establish an institutional repository is costly.

2.16.2 Inadequate power and ICT connectivity and infrastructure

Another challenge is the issue of fast and reliable internet connection as well as adequate information and communication technology infrastructure. Most African institutions have to deal with the problem of the slow speed and unreliability of internet connectivity. In the Sub Saharan Africa region, the problem of low and yet expensive internet bandwidth persists. Institutional repositories require reliable and fast internet connections as the mode of availability of materials are in PDF files. Jensen (2006:1 as quoted by Christians, 2008:32) states:

> "Bandwidth is the life-blood of the world's knowledge economy, but it is scarcest where it is most needed – in the developing nations of Africa which require low cost communications to accelerate their socio-economic development. Few schools, libraries, universities and research centres on the continent have any internet access. For those that can afford it, their costs are usually thousands of times higher than for their counterparts in the developed world, and even Africa's most well-endowed centres of excellence have

> less bandwidth than a home broadband user in North America or Europe, and it must be shared amongst hundreds or even thousands of users" (Christians, 2008:32)

African universities pay a lot of money for internet bandwidth. Christian (2008) found that 1.544 megabits per second of bandwidth, which African universities pay a lot for, is less than the one used by home consumers (internet connections for home use) in the developed countries and yet pay very little a month. For example, while home consumers in North America and Europe pay less than $100.00 a month; African universities pay more than $10, 000.00 a month for less internet bandwidth. According to Christian (2008), the major factor responsible for the high cost of internet bandwidth in Africa is that it uses satellite bandwidth and not the optic fibre infrastructure bandwidth that reduces costs and is used in developed countries. From the study conducted by Jensen (2006), he found that most of Africa is still not connected to the global fibre infrastructure. He concluded that optic fibre is the only way to supply sufficient international low-cost bandwidth which Africa needs (Jensen, 2006:3).

There is has been tremendous developments in Africa with regard to internet connectivity. A number of initiatives have been in place and some have been successful while other projects are still under way. Song (2012) has noted a number of fibre optic undersea cables connecting Africa. These are: the SEACOM project that was carried out in 2009 and funded by private companies. It directly connects South Africa and East African countries; it passes through India and Europe covering a distance of 17 000 km. The government of Kenya also funded a project called TEAMS (The Lower Indian Ocean Network) that connects Madagascar and Mauritius that was also completed in 2009. Another project in East Africa was the East African Submarine Cable system (EASSy) connected East Africa to the rest of the world. The project was estimated to cost a total of US$ 200 million, and its implementation was expected to start early 2006 and be operational by mid 2007 but finished in 2010. While the ESSAy project was underway, another project the New Partnership for Africa's Development (NEPAD) that linked landlocked countries to the the internet from the ESSAy project was also carried out. In 2010, another project the SAT-3/WASC (South Africa Trans-Atlantic - West Africa Submarine Cable) connects Portugal to South Africa, which also connects Spain to Europe goes through a number of West Africa countries was also completed (Dhliwayo, 2004).

The latest developments for 2011 and 2012 are the West African Cable System (WACS) project in which a submarine communication cable links South Africa and the UK along with some West African countries in 2011. The internet communication system was expected to start working by end of 2011(Dhliwayo, 2004). Lastly, there is a project that is to be carried out in 2012 called the African Coast Europe (ACE) project where there is a cable, which is planned to go along the west coast of Africa passing between France and South Africa (Song, 2012).

All the above projects are a sign that submarine fibre optic cables now surround the whole of Africa and many of its countries are now connected to the internet and the world. Full bandwidth problems will soon disappear because of the extensive coverage and networks capabilities.

Another problem closely associated with internet connectivity and bandwidth is that of electricity power supply. An institutional repository needs a stable and regular electricity supply for the ICT hardware (i.e. servers, computers and scanners) to function.

2.16.3 Copyright and intellectual property rights issues

Another challenge for institutional repositories is that of copyright. Many authors and researchers are essentially ignorant about their intellectual property rights and do not know that they can negotiate with publishers about depositing their materials on open access institutional repositories. Most of the time they sign away their rights and transfer copyright from the author to the publisher instead of negotiating for creative commons licenses (Moahi, 2009). This has been discussed in section 2.13 above. Institutional repositories are negatively affected as authors believe they are not able to deposit their work in them. A lack of deposits causes institutional repositories to have low contents figures, which is the core of any successful institutional repository.

2.16.4 Lack of advocacy

Christian (2008) observed that ignorance of the open access institutional repository seems to be one of the challenges to the development of open access institutional repositories in developing countries, especially African countries. Advocacy is the only way through which these stakeholders can find out about the concepts of open access and understanding what it is.

Christian (2008) argues that if institutional repositories are established without accompanying advocacy on open access, the result will be that academic authors will not appreciate and understand the impact of depositing in institutional repositories on their academic profile and impact. There is a great need to inform and educate academics and researchers about the role and benefits of institutional repositories and how they may advance their visibility.

2.16.5 Lack of faculty involvement leads to low content

Lack of faculty involvement in matters about institutional repository is one of the major challenges that implementers face as it leads to low contents figures in the institutional repository. Problems surrounding this challenge that have been identified include the reluctance of authors to self-archive because of fear of plagiarism and having their ideas stolen, confusion whether depositing one's work is publishing, and the conception that depositing work in institutional repositories is time consuming (Moahi, 2009). A successful institutional repository depends on the willingness of authors to deposit their work voluntarily. This contributes to difficulties in generating content for institutional repository (Jain, et al., 2009).

2.16.6 Technological factors

Institutions must be aware that technology becomes obsolete very quickly and must bear in mind the costs that will be associated with long-term preservation of research output in order to sustain the institutional repository (cf. 2.11.)

2.16.7 Institutional culture and politics

Institutional repositories need the support and commitment from the management and academic staff in the institution, but sometimes this is very difficult. Lynch explained by stating, "stewardship is easy and inexpensive to claim; it is expensive and difficult to honour, and perhaps it will prove to be all too easy to later abdicate" (2003: 334). This means that whenever a library in any institution thinks of establishing an institutional repository, it must seriously consider who will support it financially as it may fail due to a of lack of funding if it is not properly managed (Jain et al, 2009).

2.17 Institutional repositories and the future

From the study that Anbu (2006) conducted, he found that the UK House of Commons, Science and Technology Committee report urged the government to allocate funds to universities to start institutional repositories and ensure long-term preservation of digital scholarship emanating from these digital repositories. In the USA, the U.S. House of Representatives Committee on Appropriations recommended that all NIH funded research should be archived in PubMed Central six months after their publication. In June 2006, the Research Council of UK (RCUK) issued the long awaited open access policy directing all its research to be published in Open Access repositories (Anbu, 2006). In 2006, the US Senate passed the Federal Research Public Access Act of 2006 on Open Access to research literature. In 2007, in Australian the government's Productivity Commission produced a report with recommendations for open access dissemination of research information to science and innovation. The report emphasised the academic, social and economic benefits brought about by open access to the country (Gray, 2007). Similar proposals were being discussed in various countries like Canada, Scotland, Norway and India. Africa is still behind other continents with regard to issues about Free and Open Source Software (FOSS). The only African country that is active on issues of FOSS is South Africa. In 2007, Cabinet approved a policy and strategy for the implementation of free and open source software (FOSS) in government.

In South Africa talks about FOSS started in 2001, a number of stakeholders and organisations carried out research on the feasibility of shifting from the use of proprietary software to FOSS in South Africa. These stakeholders started to talk to government about the shift and government opened debates and commissioned a number of departments to look into the issue. The National Advisory Committee on Innovation (NACI) for the Government Information Technology Officers Council (GITOC) conducted a study, which found out that the government was spending a lot on proprietary software. In 2002 and 2003, Cabinet adopted policy recommendations from the Government IT Officer's Council (GITOC) pertaining to Free & Open Source Software FOSS. Finally, in 2007 Cabinet approved a policy on FOSS. Below are the five points from the policy:

1) "The South African Government will implement FOSS unless proprietary software is demonstrated to be significantly superior. Whenever the advantages of FOSS and

proprietary software are comparable, FOSS will be implemented when choosing a software solution for a new project. Whenever FOSS is not implemented, then reasons must be provided in order to justify the implementation of proprietary software.

2) The South African Government will migrate current proprietary software to FOSS whenever comparable software exists.

3) All new software developed for or by the South African Government will be based on open standards, adherent to FOSS principles, and licensed using a FOSS license where possible.

4) The South African Government will ensure all Government content and content developed using Government resources is made Open Content, unless analysis on specific content shows that proprietary licensing or confidentiality is substantially beneficial.

5) The South African Government will encourage the use of Open Content and Open Standards within South Africa." (Department of Public Service & Administration, 2006:2)

The South African Research Information Services (SARIS) was formed in conjunction with the Academy of Science of South Africa (ASSAf) to lead the open access movement by encouraging institutions to establish institutional repositories (Van Deventer and Pienaar, 2008). In 2002, a national research strategy was published as a result of the SARIS project which was carried out. The Department of Science and Technology (DST) and the Department of Higher Education and Training (DHET) funded Academy of Science of South Africa (ASSAf) to establish the Scientific Electronic Library Online (SciELO) South Africa. This is an open access database for online journals mainly for researchers. Another example of open access initiatives in South Africa is the Human Sciences Research Council (HSRC), which offers both open access publishing and printing publishing. South Africa is so far the largest producer of research articles in Africa and many of its universities have established institutional repositories to promote the open access movement (SciELO, 2012).

Lynch (2003) states that in future, higher education institutions will not only run institutional repositories but will also form institutional repository clusters to offer services to their communities. This could mean that institutional repositories would be a basic need in the development of regional or disciplinary consortia among universities or libraries. Lynch (2003: 336) also proposed "federating" institutional repositories. This refers to "cross-repository

searches, swaps of storage between institutional repositories to gain geographic and systems diversity in pursuit of backup, preservation, and disaster recovery or other capabilities are made possible" Lynch (2003: 336). Lynch (2003) further argues that this will be a fruitful area for exploration and innovation. Federation of institutional repositories may also include the development of arrangements that recognize and facilitate movement of faculty researchers and cross-institutional collaborations. Webley et al. (2011), Thuku, (2011) and Ackermann (2011) reported at the the ETD 2011 conference, that many countries around the world have started creating National Electronic Thesis and Dissertation Portals. For example, in Africa there is the Database of African Theses and Dissertations (DATAD) (Webley et al., 2011). This is almost similar to what Lynch has forecasted.

It is foreseen that public libraries might join forces with local government, local historical societies, local museums and archives, and members of their local communities to establish community repositories (Lynch, 2003). With all these forecasts, one concludes that establishing an institutional repository may yield numerous benefits for communities and surroundings.

Lynch (2003) also sees the future for repositories moving toward publishing services and the services libraries can offer to faculty. The forecast is that the next generation of institutional repositories must focus on providing specific services that have value, which in turn are marketed to faculty members and are supported by the university or institutions' administration. These services include librarians actively soliciting and depositing materials for faculty, thus increasing their authors' web profiles.

2.18 Conclusion

The chapter has discussed issues regarding the establishment of institutional repositories. It started with an exploration of the historical background of institutional repositories and the reasons that prompted their development. It also traced international policy and principle statements about the open access movement. The international initiatives on institutional repositories in Europe, America, Canada, and Australia were discussed and showing how all these initiatives shaped events that later created possibilities for establishing institutional repositories in Africa. The main points considered are that the establishment of institutional repositories seems to create similar challenges irrespective of the country or continent in which

they are established. Obviously, where funding is available, matters are made easier. This was the case for USA and Europe, while Africa is still grappling with issues of poor electricity supply, information and communication technology infrastructure and internet bandwidth. Although Africa is fighting to overcome its challenges, there are other countries in Africa, namely South Africa that are managing and showing some progress.

The chapter that follows discusses the methodology chosen to carry out the research project and the strategies that will be used for collecting data.

CHAPTER 3

RESEARCH METHODOLOGY AND DESIGN

3.1 Introduction

As noted in Chapter one, the purpose of this study is to explore the development, establishment and implementation of a digital institutional repository in the University of Cape Town Libraries (UCT Libraries). This chapter discusses the research design and methodology for the study. The research instruments and data collecting techniques will be discussed.

3.2 Primary Research questions

The methodology chosen for this study on the development, establishment and implementation of an institutional repository considers the research questions:

- What are the steps involved in the development, establishment and implementation of an institutional repository and what measures have to be taken to overcome identified challenges?
- How does the UCT institutional repository compare and contrast with that at UP?
- To what extent will the process followed at UCT in developing an institutional repository lead to a resource that meets with criteria for success?

3.2.1 Secondary research Questions

- What are the factors that influenced UCT to establish an institutional repository?
- Which resources and structures were put in place in developing the institutional repository at UCT?
- To what extent is the repository an open access initiative?
- Are the existing departmental digital repositories being brought together under the UCT Institutional Repository?
- What criteria that could be used to evaluate the success of an IR?

- What challenges have been encountered in the establishment of the institutional repository?
- How are these challenges overcome?
- Compare the UCT institutional repository with that at UP.
- Has the UCT been able to meet the criteria used to evaluate the success of an institutional repository?

3.3 Qualitative Research Methodology

This research project follows a qualitative research methodology. Qualitative research deals with the real world, human beings and their experiences, social context and multiple other realities (Bogdan, 1992). It provides a method to probe deeply into the inner world of persons and reconstruct their thought processes and lifestyles (Nachmias & Nachmias, 1996). The emphasis is on the natural context of the environment within which an activity takes place. It is far more situation specific, provides a contextualised description of situations, and allows phenomena to be studied in detail (Mouton and Marais, 1990). Campbell-Meier (2008) defines qualitative research as:

> "an inquiry process of understanding based on distinct methodological traditions of inquiry that explore a social or human problem. The researcher builds a complex, holistic picture, analyzes words, reports detailed views of informants, and conducts the study in a natural setting " (Campbell-Meier, 2008: 33)

Researchers such as Mouton and Marais, 1990; Neuman, 1994; Nachmias and Nachmias, 1996; and Babbie and Mouton, 2001, assert that qualitative research, in the broadest sense, is the research that produces descriptive data, people's own written or spoken words and observable behaviour. Quantitative research, on the other hand, tends to be more formalised and follows the positivist approach in which the researcher works with clearly defined concepts and tries to be as objective as possible, and operates at a distance. In qualitative research, the researcher becomes part of the research (Mouton and Marais, 1990).

The most important difference between qualitative and quantitative research is the way in which each approach treats and collects data (Babbie and Mouton, 2001). Data collecting techniques may appear similar, but the orientation differs considerably. Quantitative methods are oriented towards following a far more structured and unobtrusive approach and obtaining quantifiable data that may be analysed using sophisticated statistical methods. Qualitative data is usually collected from individuals or small groups. The researcher is involved with the instrument and the information generated is often not numerically analysed or presented (Babbie and Mouton, 2001). Some of the characteristics of qualitative research are listed below:

- Provides an in-depth understanding of social and human problem and experience of a selected topic,
- Enhances understanding of social processes,
- Is conducted in a natural setting,
- Emphasises the use of human beings as research instruments,
- Non experimental and does not require dependent and independent variables,
- Uses inductive methods and provide descriptive research results,
- Sampling is not predetermined, and
- It attempts to include as many participants as possible (Babbie and Mouton, 2001).

Although qualitative research methods are regarded as useful for the in-depth understanding of social phenomena, they are criticized because they deal with small sample sizes and therefore, are incapable of generalisations. They are also said to be subjective and biased as the researcher is more involved with the research process and the instruments. There are also expensive and time-consuming factors and the research has little control over the results of the study.

For this research, the case study method was chosen over other investigative methods because of its appropriateness for the context of the UCT as an institution. It falls within the qualitative research methodologies, and has allowed the researcher to gain an in-depth understanding of the topic from the participants who were interviewed.

3.3.1 The Case Study

Babbie and Mouton (2001) state that case studies are best at depicting institutional and organisational best practices, policy implementation or evaluation. This study follows developments towards the implementation of a policy to establish an institutional repository and this is why this method was seen as appropriate.

According Yin (1994) the case study research method is

> "an empirical inquiry that investigates a contemporary phenomenon within its real-life context; when the boundaries between phenomenon and context are not clearly evident; and in which multiple sources of evidence are used." (Yin, 1994: 23)

A case study thus falls within the qualitative research approaches. When one uses a case study, it means that the researcher is not intending to study the entire organization. Instead the intention is the collection and detailed presentation of what one wants to explore. Case studies are intended to focus on a particular issue, a particular participant or a small group feature or unit of analysis, frequently including the accounts of subjects themselves (Umit, 2005). It draws conclusions only about that participant or group and only in that specific context (Umit, 2005; Yin, 2009; and Campbel-Meier, 2008). Additionally Noor (2008) states that case studies are also useful where one needs to understand some particular problem or situation in great depth, and where one can identify cases that are rich in information.

"Case studies can be either single or multiple-case designs" (Yin, 2003: 47). Single cases are used to confirm or challenge a theory or to represent a unique or extreme case (Yin, 2003). On the one hand, single-case studies are important in showing hidden matters that can only be accessed by people inside the organisation. Researchers using this type of case study need to do a careful and thorough investigation, in order to avoid misinterpreting issues. On the other hand multiple-case studies follow replication logic, thus having to repeat cases (Campbel-Meier, 2008: 29). According to Campbel-Meier (2008: 31) "each individual case study in the multiple-case consists of a whole study", in which facts are gathered from various sources and conclusions drawn from those facts (Campbel-Meier, 2008).

This research employs the single-case design. It is a case study of a multi-dimensional academic institution, namely UCT thus a case study which looks at an organisation. Case studies are recommended in the exploratory research of single units, they are also best for depicting institutional and organisational best practices (Babbie and Mouton, 2001).

3.3.1.1 Arguments against the use of case studies

Some researchers do not approve of the case study, arguing that since case studies deal with smaller populations or a small number of cases, they offer no grounds for establishing reliability or generalizability of findings and conclusions (Campbel-Meier, 2008). Yin (2009) pointed out that the generalization of results, from either single or multiple designs of case studies can be made into theory and not extended to populations.

Case study methodology is said to be 'microscopic,' meaning that it is a thorough study that looks at very few cases (Yin, 2003). Researchers who are against case study methodology argue that case study research lacks quantitative measurements in their research design and analysis (Campbel-Meier, 2008). Some of the researchers dismiss case study research as useful only as an exploratory tool (Yin, 2003).

Even with these conflicting ideas about case studies, researchers continue to use the case study research method with success in carefully planned and crafted studies of real-life situations, issues, and problems. Reports on case studies from many disciplines are widely available in the literature (Noor, 2008).

From the arguments above, it is clear that case studies need to be carefully conducted so that the researcher does not fall into the trap of being too involved with the subjects under study. In so doing, bias is avoided. One needs to employ the strategies outlined by Yin (2009) which are:

- Asking good questions,
- Being a good listener,
- Being adaptive and flexible,
- Having a firm grasp of issues being studied, and
- Being unbiased by preconceived notions (Yin, 2009: 69-72).

Case studies produce a clear and in-depth understanding of a situation. They are valuable method of research, with distinctive characteristics that make it ideal for many types of investigations. It can also be used in combination with other methods such as surveys and experiments. A case study approach was chosen in order to understand the reasons why UCT is establishing an institutional repository and how it will be developed and implemented.

3.3.1.2 Types of case studies

According to Babbie and Mouton (2001), case studies are of six types;

a) individual cases,

b) community cases,

c) social group studies,

d) studies of organisations and institutions,

e) studies of events, roles, and relationships and

f) studies of countries and nations (2001: 281).

For the current research, a case study of organisations and institutions will be used to investigate the development and implementation of an institutional repository. This is a single unit case study of a multi-dimensional academic institution, UCT.

3.3.1.3 Reliability and Validity of the study

According to Yin (2003) the problem of construct validity in case study research, namely the extent to which a measurement corresponds to the concepts under study, is problematic. Therefore, a researcher might use multiple sources of evidence to strengthen internal validity in descriptive or exploratory case study research. Case studies are strong in reliability because of the use of multiple sources of evidence.

Another way to increase validity is by giving the case study report to the key informants to review and correct any mistakes they may come across. Here the objective is to be sure that if

another researcher followed the same procedures, he/she would arrive at the same conclusions and findings. According to Yin (2009), reducing the errors and biases in a case study can be achieved in many ways. Some of the ways noted by Yin (2009) are, when a researcher writes down all the steps that were taken when following a case, this is called the case study protocol. The second way of reducing errors and bias is when a researcher develops a case study database. The goal of reliability is to minimize the errors and biases in a study.

3.4 Research Design

The research design is the blueprint that assists the researcher in solving the problems pertaining to the investigation. According to Babbie & Mouton (2001) it consists of the "set of decisions regarding the topic to be studied in the population, outlining which research methods and for what purpose is the research carried out" (2001: 104). According to Nachmias & Nachmias (1996), a research design is a plan that,

> "...guides the investigator in the process of collecting, analysing, and interpreting observations. It is a logical model of proof that allows the researcher to draw inferences concerning causal relations among variables under investigation" (1996: 77-78).

Kerlinger (1964) defines a research design as "the plan, structure and strategy of investigation conceived to obtain answers to research questions and to control variance." Babbie and Mouton (2001) define a research design as the "plan or blueprint on how you intend conducting the research" (2001:74). When one is designing research, the methods or procedures used must be investigated, namely the types of measurement, sampling, data collection and data analysis methods. One must ask what questions are going to be studied, what data are relevant, what data to collect, and how to analyse the results from the collected data (Babbie & Mouton, 2001). The research design helps one to avoid a situation in which the evidence does not address the initial research question (Yin, 2009). The components of a case study research design, as outlined by Yin (2009) are, the study question, its proposition, its unit of analysis, the logic linking that data to the proposition and the criteria for interpreting the findings (Yin, 2009: 27), i.e. the development of 'how' and 'why' questions and their definition.

The first task of the researcher is that of formulating the research question and sub questions and capturing what one is interested in studying. Case studies are seen as being concerned with 'how' and 'why' things happen, they allow the investigation of contextual realities and the differences between what was planned and what actually occurred (Yin, 2003; Umit, 2005; Noor, 2008; and Campbel-Meier, 2008).

An exploratory study like this, which has a stated criterion according to which success will be judged, will be discussed in detail below. The unit of analysis could be an individual, an event or process. The logic linking the data to the proposition is the specification of methods and procedures for acquiring the information needed for solving the problem. The data collecting techniques such as questionnaires, interviews, and focus groups are decided while framing the research design. Finally, the criteria for interpreting the findings enable the researcher to arrive at a certain meaningful conclusion at the end of the proposed study and thus the research design (Yin, 2009). The unit of analysis for this research project is the organisation.

When Babbie and Mouton (2001) discussed case studies, they mentioned four general design principles, which are conceptualisation, contextual detail, multiple sources of data and conceptual categorisation. These are discussed below with reference to this research study.

Conceptualisation will be done through the careful description and definition of the concept of the institutional repository and the open access model of publishing, and a formulation of the research question and sub-questions. Contextual detail is given through the report of the historical background of the departmental digital repositories, the first digitisation project by the Manuscripts and Archives Department of UCT Libraries in 2001, the opening of the Digital Initiatives Unit in the UCT Libraries and the move towards the development of the Institutional Repository at UCT. Multiple sources of data such as documents, participant observation, interviews and physical artefacts will be explored. Conceptual categorisation is the process in which ideas and concepts are recognized, differentiated and understood. For this research project this will be done through grouping into categories, the themes and topics deduced from the interview and discussed to bring out the main arguments and specific points of the research topic and question.

3.5 Population of the study

A number of major researchers in the field of case study research such as Tellis, (1997), Yin, (2009), and Stake, (1995) all assert that, "case study research is not sampling research" (Tellis, 1997). However, they all agree that when selecting cases this must be done to maximize what can be learned, in the time available for the study. The other important issue is that the unit of analysis is a critical factor in the case study as "it is typically a system of action rather than an individual or group of individuals" (Tellis, 1997). This is because case studies tend to be selective, focusing on one or two issues that are fundamental to understanding the system being examined and because of that, there is no sampling of population required. The research therefore selects the population under study, based on its knowledge about the issues being studied.

The target population of this study is informed and knowledgeable UCT academic and library staff members on matters pertaining to the development and establishment of an Institutional Repository at UCT. Academic and library staff members will be interviewed from the departments which already have departmental repositories, namely the

- Faculty of Law,
- Department of Computer Science,
- Manuscripts and Archives Department (UCT Libraries),
- Open Educational Resources – UCT - Centre for Educational Technology, and
- UCT Main Library staff that are involved in the institutional repository development.

Members of these departments will be asked for their views and perceptions about their respective repositories. Academic staff will be asked to give their view on institutional repositories and what they know about open access. Library staff will give a detailed explanation of how the library started the repository, from the time it was proposed, its development, implementation, and the challenges and solutions that have been identified.

3.6 Data Collection Techniques

When conducting a case study, there are a number of data collecting techniques that can be used, and there are at least six sources of evidence in case studies, as stated by Yin (2009).These include documents, archival records, direct observation, participant-observation, interviews, and physical artifacts.

Although there are many ways of collecting data for case studies, no single source has a complete advantage over the others, but they might be complementary and may be used in collaboration (Yin, 2009). Thus, a case study should use as many sources as are relevant to the study. Tellis (1997) is of the view that doing a case study depends on using multiple techniques as it develops a rich store of information, permitting the analysis of a single issue from several perspectives. This is called "triangulating the data." There are different types of triangulation: data triangulation, investigator triangulation, theory and methodological triangulation (Neuman, 1994). Data triangulation will be used in this study as multiple sources of evidence will be used.

The research project will thus employ a triangulated approach, which is recommended when the limitations of the other data collecting techniques can be complemented by one another. Triangulation is generally considered to be the best way to enhance validity and reliability in qualitative research. It can therefore be summarised that "triangulation is important because it compensates for individual shortcomings of each method. It ensures flexibility, systematisation and serves as a validity check." (Babbie & Mouton, 2001: 310). It affords the maximum utilisation and analysis of data.

Below is a discussion of the sources of evidence that will be used to collect the data for the case study and how they will be used.

3.6.1 Documents

Documents used during a case study could be letters, memoranda, and personal documents such as diaries, agendas, minutes of meetings, written reports, administrative documents, newspaper articles, news clippings, or any documents that are related to the investigation. The documents serve to strengthen the evidence from other sources while they are also useful for making

inferences about events. According to Yin (2009), documentary information is relevant to every case study. Although documents may be biased or lack accuracy, they are useful.

Documents can lead to false leads in the hands of inexperienced researchers, which have been a criticism of case study research (Umit, 2005). When dealing with documents, it is important to review them carefully to avoid incorrect data being included in the database. This establishes the validity of the documents. Yin (1994) warns inexperienced researchers not to rely too much on documents as evidence, as there could be the danger of mistaking the truthfulness and authenticity of some documents.

In this research, the following types of documents will be investigated:

- minutes of meetings,
- administrative documents or proposals for the establishment of the institutional repository,
- written reports related to the establishment of institutional repository,
- progress reports,
- inventory reports,
- letters written to request funding or permission to establish an institutional repository and
- news clippings and other materials from the mass media or the university newspaper.

These documents will be from the year 2009 when the establishment of the unit was proposed, to the present time.

3.6.2 Archival Records

Archival records are useful in some studies since they could include service records, maps, charts, lists of names, survey data, organizational records such as budget records, and public files. The investigator has to be extremely careful about details in determining the origin of the records and their accuracy (Tellis, 1997). An important statement by Yin (2009) is that most archival records were produced for a specific purpose and for a specific audience other than the case study investigation. These conditions must be fully appreciated in interpreting the usefulness and accuracy of the records. For this study, archival records will not be used.

3.6.3 Direct Observation

Direct observation refers to when the researcher observes what the research subjects (observing activities) are doing or record behaviours by the subjects in order to get first-hand information about the topic under study. Direct observation is an important data collection technique as it gives the researcher additional information about the subjects under study or the topic that is being studied. In a situation where there are more observers, it helps enhance the reliability of the study (Umit, 2005). Instead of direct observation, this study will use participant observation as discussed below.

3.6.4 Participant Observation

Participant observation is a type of observation where the researcher does not merely watch what is happening and observe events but becomes part of what is happening. One becomes involved in what is being done in the organisation. For example, being one of the staff members in the workplace, serving customers, attending meetings, organising events as well as taking part in those events and taking decisions (Yin, 2009).

This technique is very commonly used by researchers in anthropological studies of different cultural or social groups. It can also be used in everyday settings such as small informal groups and large organisations (Yin, 2009). The advantage of using this technique is that it provides the researcher with opportunities to obtain inside information from the organisation that one would not achieve by merely interviewing people or direct observation. The second advantage, as stated by Yin (1994) is that the researcher has the ability to perceive reality from the viewpoint of someone "inside" rather than external to it. Another advantage is that the researcher is able to manipulate some situations such as calling short meetings of some of the participants in the case study.

While this method is good for getting inside information from the organisation under study, it has some disadvantages that researchers should carefully consider. The main disadvantage of this technique is that a researcher, as an active participant may become biased. The second disadvantage is that the researcher may end up supporting the organisation being studied (Yin, 2003). The third disadvantage is that the researcher can become so involved with the work being

done in the organisation that limited or no time remains to ask questions about the phenomenon under study or to document notes as a direct observer. Lastly, if the organisation under study is rather large and has branches in various places, it will not allow the researcher to participate and observe important events in the various places, thereby limiting their scope.

For this research, direct-participation will be used. The researcher requested to be briefly attached to the Digital Initiatives Unit to assist in carrying out duties that the staff members perform on a day-to-day basis. This allowed the researcher to understand how the Institutional Repository at UCT was being developed. A diary of the visits to the Digital Initiatives Unit is kept. The participant observation ran from 1st to the 30th of June 2011. Thereafter, interviews with key informants from the UCT Chancellor Oppenheimer Main Library were conducted from the 3rd to the 14th of July 2011. The analysis of the participation observation, evidence found from scrutinising documents, physical artefacts, reports and interviews, are discussed in chapter Four (cf 4.5).

3.6.5 Interviews

There are different types of interviews that a researcher can use when doing a case study. These include open-ended in-depth interviews, focused interviews, and structured or survey interviews. Interviews have been identified as the most important information sources for researchers doing case study. Key interviewees are normally asked to elaborate on issues and events and also give their opinions in open ended in-depth interviews (Tellis, 1997). These key informants may sometimes support information from other sources. Researchers who use interviews are warned not to depend too much on key informants (Tellis, 1997). Tellis (1997) proposes that "the researcher must seek the same data from other sources to verify its authenticity or seek the same data from other informants who may have been working closely with the informant" (Tellis, 1997).

The focused interview is "designed to determine the responses of persons exposed to a situation" (Merton and Kendall, 1946: 541). It can also be used to draw upon respondents' attitudes, feelings, beliefs, experiences and reactions in a way in which would not be feasible using other methods, such as observation, one-to-one interviewing, or questionnaire surveys (Gibbs, 1997).

People who participate in this kind of interview are normally experts in their fields and have a lot of information on the topic under investigation. The researcher would also need to set specific questions for the panel to get different perspectives about the topic. The technique of the focused interview is often used to confirm data collected from another source (Tellis, 1997).

A further type of interview is the semi-structured interview. This is where a researcher sets formal question in advance in the form of a questionnaire where respondents will have to fill in the questionnaire. After administering a questionnaire, the researcher may choose to have interviews with respondents as follow up to clarify answers from the questionnaire (Tellis, 1997).

For this research, semi-structured interviews with selected librarians were held. Questions were derived from the information taken from other data collecting instruments such as documents, direct observation and physical artefacts. Semi-structured interviews were based on guiding questions. These were face-to-face interviews so that the interviewees could give the researcher in-depth information and deeper insight. It also allows for the establishment of rapport. The researcher used open-ended in-depth interviews, with key respondents commenting about certain events in the development of the UCT Institutional Repository. They proposed solutions or provided insight into events that have taken place. The interview took into consideration the type, structure, sequence and wording of the questions asked as these aspects could have affected the respondents' motivation.

Two sets of interviews were conducted. The first set of interviews were conducted in March 2011. These were exploratory interviews which allowed the researcher to gain a better understanding of the state of departmental repositories at UCT. These interviews took place at the beginning of the year and were unstructured interviews. They consisted of a list of prepared questions that were derived from the literature.

The second set of explanatory interviews took place between the 1st and the 14th of July. These interviews became the core sets of interviews of this research study, as they were held with key informants who had pertinent knowledge to particular questions. The semi-structured interviews were based on insights gained during the researcher's participant observation work, information found in the documents collected, and on physical artefacts records. Furthermore, the nature of

semi-structured interviews allowed the researcher to ask additional questions based on the responses of the interviewees.

3.6.6 Physical Artefacts

Physical artefacts could be any physical or cultural artefact, tools, notebooks, computer output, instruments, a technological device, a work of art, or some other physical evidence that may be collected during the study as part of a field visit (Umit, 2005). The researcher's perspective can be broadened by the discovery of the information contained both from and within these physical objects or artefacts.

In this study, list of the artefacts in the Manuscripts and Archives Department was taken and records of materials found in the Digitisation unit were investigated. These included records of photos, images, and maps. At the digitisation unit, technological devices such as scanners, printers and the computers may have some information about the establishment of the institutional repository. The study will determine the number of digitised artefacts and review how the artefacts are stored for preservation.

3.6.7 Three Guidelines for Case Study Data Collection

There are three guidelines of data collection that have been suggested by Yin (2009). These include using multiple sources of data, creating a case study database and maintaining a chain of evidence (Yin, 2009). For this research project, the three guidelines for data collection are employed. The reason for using various sources of information allows one to triangulate evidence, as mentioned above. This was applied in the study by using documents, interviews, participation observation and physical artefacts. This will increase construct validity and allow for the triangulation of data.

The second guideline requires that the researcher should create a study database. One should first create the database for the data, and secondly a database for the reports that one will be writing. The databases help a researcher in organizing and providing proof that the data have been collected for the case study. The researcher should make the databases simple, and include relevant information that will be used and understood by other researchers. For this research project, a database was created containing case study notes which were written clearly and made

available for later use. These notes according to Yin (2009), may be in the form of interview results, observations, or document analysis. The notes can either be typed, handwritten or on audio tapes. Secondly, case study documents that are relevant to the case study and separate from the final report will be included. In short, a case study database increases the reliability of the case study (Yin, 2009).

Lastly, a researcher should collect a series of facts to increase the reliability of the study. Yin (1994) proposes that one creates citations to the case study database where the real evidence or facts are located. In this research, it will be ensured that no original evidence is lost in the process of collecting data either because of carelessness or through bias. The process of presenting the data should be straightforward and indicate that all the evidence presented in the case study is undoubtedly the same as that collected during the data collection process (Yin, 2009: 122). The original evidence is stored in the database allowing it to be consulted for cross-reference checks.

3.7 Data Analysis

Data analysis is defined as "the process of systematically and physically arranging interview transcripts, field notes and other collected data that was accumulated during data collection" (Bogdan and Taylor, 1975: 153). Data analysis includes working with data, organising and breaking it into manageable pieces, searching for patterns, discovering what is important and deciding what to tell others (Bogdan and Taylor, 1975). It provides a way of examining and interpreting data. In qualitative research, data analysis is inductive in nature, because critical themes and meaning emerge from the data requiring creativity and the necessary skills to communicate and interpret the data (Smith, 2006).

Data analysis of case study methodology is the most difficult as it is not conducive to statistical analysis and as such the process of analysing case study textual data is not easy (Umit, 2005). Some strategies that could be used generally when dealing with case study data, were proposed by Yin (2009).

These consist of four techniques, namely theoretical propositions, case study descriptions, using both qualitative and quantitative data, and examining rival explanations. When one is using theoretical propositions it means the researcher is guided by theoretical orientations or propositions that would have helped the researcher to formulate the design to the case study. The design of the case study includes the objectives and the research questions of the case study. It helps the researcher to focus attention on certain data and ignore other data. Case study descriptions help the researcher to describe the framework for organising the case study normally based on the general characteristics and relations of the phenomena in question (Yin, 2009). Both qualitative and quantitative data is used when the case study also has a large amount of quantitative data that will be analysed with statistical techniques. The last technique of examining rival explanations occurs when the researcher tries to define and test these differing rival explanations (Yin, 2009). This can be used with all the three of the techniques already mentioned above.

For this project, data analysis techniques that will be used are theoretical propositions and case study descriptions. These will be supported by data from the interviews.

3.8 Conclusion

This chapter has presented the research methodology adopted for this study. The major conceptual responsibilities of the qualitative case study researcher are listed below.

- bounding the case, i.e conceptualizing the object of study,
- selecting the phenomena
- identifying themes or issues,
- seeking patterns of data to develop issues,
- triangulating key observations and bases of interpretation,
- selecting alternative interpretations to pursue, and
- developing assertions or generalizations about the case (Stake, 1995).

The research followed the qualitative research methodology as it is viewed as the most appropriate method for social research studies. Library and Information Science falls within the broad domain of the Social Sciences and is concerned with the interaction between people and information hence a social system. Mouton and Marais (1990) states that "Social Science research is a collaborative human activity in which social reality is studied objectively with the aim of gaining a valid understanding of it" (1990: 7). Within the qualitative research methodology, a case study approach was chosen. This was based on its strength to provide a better understanding of the context and process surrounding the development of UCT's institutional repository. Several techniques of data collection were used to increase the validity of the findings.

The next chapter will focus on a case study description of the establishment of the UCT institutional repository, how it started, who assisted with funding, the challenges encountered and how these were overcome.

CHAPTER 4

ESTABLISHING AN INSTUTIONAL REPOSITORY AT THE UNIVERSITY OF CAPE TOWN: A CASE STUDY

4.1 Introduction

In this chapter, the discussion will focus on a case study of the UCT Institutional Repository and how it was developed and established. It starts providing an overview of the repository and then discusses the contents of the repository. The chapter will also discuss the approaches and strategies employed by the researcher to analyse collected data. It reports on the data collection within the following four parameters, i.e. participation observation, administrative documents, physical artefacts and, finally, it discusses interviews with informed experts in the field. The challenges encountered and the results that emerged from the data analysis process are also presented.

4.2 Towards developing an institutional repository at the UCT

As noted in section 1.3.1, the Manuscripts and Archives Department of the UCT Libraries began digitizing selected material in 2001 (Dunlop and Hart, 2005). It was a small-scale pilot project, based on the unique collection of records about the San people. As mentioned earlier, other digitization projects in UCT were instigated by individual departments. In 2003, the Computer Science started to digitize the theses and dissertations of the students in their department. Similarly, the LawSpace Digital Departmental Repository was started in 2005, with the aim of digitizing theses and dissertations of the students in the Faculty of Law. At present, there are plans to combine all the departmental repositories under the new UCT institutional repository which will be housed within the UCT Libraries.

4.3 How the Institutional Repository was established in the UCT Libraries

Interviews were carried out to understand why the first digital initiatives at the UCT Libraries were conducted at a relatively small scale in the early 2000s. Two reasons that were identified were cost and staff resources. Firstly, it was realized that establishing an institutional repository needed a budget for staffing, as well as hardware and software in order to start up the project. Secondly, appropriately trained members of staff were required.

From around 2006, the Executive Director of UCT Libraries repeatedly requested the University to budget for the implementation of an institutional repository. The deans and other members of the executive were supportive, but according to the Executive Director, in a highly competitive fiscal environment, this request always came closely behind higher priorities. In September 2009, the Executive Director, submitted the UCT Libraries' strategic plan and priorities for years 2010-2014 to the University Senate, outlining in detail the plans of the libraries for the next five years (Rapp, 2009). This document explains the impact that the Carnegie Grant had on the establishment of the UCT Institutional Repository.

In 2009, the Carnegie Corporation of New York awarded the Research Libraries Consortium, of which UCT is a member (including Witwatersrand University and University of Kwa-Zulu Natal), a grant amount of $2.5 million. This was for a three-year project directed at supporting research and library staff development at these institutions (Rapp, 2009). Among other projects, this grant provided funds to develop a new Digital Initiatives Unit with up-to-date equipment and expert staffing (Rapp, 2009). Other projects involved creating a research wing for postgraduates, the Research Commons, as well as software for sophisticated search and resource discovery tools for a digital information portal. The Unit was charged with spearheading the establishment of an institutional repository, which would be operational by the end of 2011. By establishing an institutional repository, the Digitization Unit, plays an important role in achieving the goals outlined in the UCT Libraries' strategic plans. It is stated that the unit would...

> "...attract, preserve, digitize, and make available via a sophisticated web portal key African archival and other resources. Digitize these resources as a contribution to African

scholarship, as an incentive for digital collaboration on the continent, and as a showcase for UCT's research presence" (Rapp, 2009: 2).

It was acknowledged that valuable UCT archival and research materials were not easily accessible to scholars and should be exposed to the world through an open access institutional repository. At the time, access to these materials was only possible by visiting UCT Libraries. Through the establishment of an institutional repository, these archival materials would be digitized and made available on the web for scholars' worldwide. The Institutional Repository will therefore:

a) " provide a web-based presence for UCT's collections and make them available to scholars in Africa and elsewhere,
b) enhance UCT's research presence on the web, and
c) enable better preservation of the primary physical resources." (Rapp, 2009:6)

Secondly, the Carnegie funding provided UCT Libraries with the opportunity of "restructuring library services to enhance support for meeting the information needs of academics and postgraduates..." (Rapp, 2009: 6). According to the UCT Libraries' strategic plan and priorities for the year 2010- 2014, the libraries would aim to…

> "...strengthen the Special Collections and Digitization Units to attract additional heritage collections and to enhance capacity to digitize special collections and research materials. In particular, focus on providing adequate and appropriate space and added curatorial expertise in specialized materials, such as film, video, photographs, and music" (Rapp, 2009: 6).

The Carnegie funding played a major role in financing the Institutional Repository at UCT, allowing the purchase of equipment such as scanners and cameras as well as DigiTool software.

4.3.1 Contents of the UCT Institution Repository

At present, the contents of the UCT Institutional Repository consist of digital collections, finding aids and electronic theses and dissertations with a total of 4063 materials. The contents in the

repository comprise mainly of material from the old digitization projects that took place from 2001. These older projects will be combined with the new contents digitized by the Digital Initiatives Unit since its inception to date. Below is a detailed description of the UCT Institutional Repository contents.

4.3.2 Digital collections

Digital collections are comprised of collections from UCT Libraries and other departments on campus. Various formats are represented, including images, sound and video files. It has a total of 1893 materials when it was last updated on the 30th September 2011. Under these, subject-based sets of records are located as follows:

- *Under exposed* materials is a collection that sets out to highlight South African (SA) talent. The works have contributed significantly to SA heritage and history but are unseen and hence the title "underexposed." It is a comprehensive and unique digital collection of photographs, displaying different photographers' work from the twentieth and twenty-first centuries. The photographs show both the pre and post-apartheid periods in South Africa (DigiTool website). This collection features newspaper photographers such as Arthur Bolton, David Hemson, Daniel Morolong, Cedric Nunn, the US Consul, John Savage, Anne Fischer, Basil Breakey, and some lesser known photographers. According to the UCT Digitool website, the project has been "supported by Duke University Library and is part of the Photographic Unit based at the Center of Curating the Archive (CCA) in the Michaelis Fine Art School UCT Campus".

- *Varsity* consists of materials from the UCT student newspaper dating from the 1900s. The paper changed its name to *UC Tattle* between 1934 and 1940.Thereafter, it then went back to use its original name, *Varsity*, which is still in use to date.

- *Sax appeal covers* show all the annual UCT "Remember And Give" (RAG) club Magazine covers from 1933 to date.

- The *Percy Kirby collection* from 1887-1970 is a large collection of musical instruments. Prof Kirby compiled these instruments after realizing that Southern African indigenous music was disappearing and dying. These instruments were made and played by indigenous Southern African groups.

- *UCT through the years* from 1900 to the present is a collection of photographs from the 1900s of the Groote Schuur campus to the present UCT campus in 2011.

- *British concentration camps* is a collection of photos and documents from the concentration camps. The camps were formed by the British army during the Anglo Boer war in the 1900s to accommodate residents of the two Boer Republics and the Orange Free State.

- *The San* photographs by Dorothea Bleek are a collection of 310 photographs. These were taken during her many travels where she identified and recorded the San languages of Southern Africa. They date back from her first trip in 1910, when she visited the Northern Cape in South Africa. They also include trips she made in the 1920s.

- *Black Sash* contains a collection of National Conference Papers on Black Sash minutes, responses by women to Apartheid legislation, a list of Black Sash members and their addresses. It contains texts and photographs containing newspaper clippings of the Black Sash and the South African Association of University Women Archive activities. There are also articles on Women's Movement for Peace.

- *Center for Popular Memory* is an archival centre that advocates for peoples' views and encourages generational dialogue between them. It also houses research in oral history, peoples' stories, and is considered as public history. The centre's main concern is people who have been traumatised and who need to be heard. The centre is located in the Manuscripts and Archives Department at UCT. It falls under the Historical Studies Department, which teaches oral and visual history research to students and organisations. It further conducts history interviews and disseminates them through various media such

as teaching, publication and the internet. Their website can be found at the following URL: http://www.popularmemory.org.za/. At the moment, the collection has articles of various research projects, conferences papers, training modules and different stories from people.

4.3.3 Finding aids

A *Finding Aid* is "an inventory of an archival collection, and provides detailed information about the content, arrangement, and context of a collection" (Digitool website). The difference between *Digital Collections* and *Finding Aids* is that *Digital Collections* are the digital materials that were uploaded by the new Digital Initiatives Unit. These materials were provided by UCT Libraries and other departments on campus from the time of its inception to date. *Finding Aids* directs users to the collections in the Manuscripts & Archives Department at UCT Libraries (Digitool, 2011).s Finding aids had 971 records when it was last updated on the 30th September 2011. They are arranged in alphabetical order by categories. Below is a list of the Finding Aids that are available from the UCT repository.

- *South African College/UCT* is a collection of documents and minutes of the history and events about UCT from the 1900s. It showcases how the current UCT started as a "white boys" college called South African College. This was the time it did not accept female students. It later became a "white only" university because of the then apartheid government with both male and female students in attendance. The first few black students were admitted in the 1920s and it remained low until the 1980s and 90s. Its name then changed to the University of Cape Town. Today it accepts students of all races and the number of black students admitted has risen to 35 percent. The other change that has taken place is that nearly half of UCT's students are female.

- *Architectural Collections* is a collection of architectural documents and drawings of buildings in the Cape Province by different architects. It is a miscellaneous collection of literary compositions or pieces by several authors. These deal with various topics on pre historic houses and buildings of South Africa, particular to the Western Cape, covering the need to preserve them. Letters, articles from newspapers, journals, reports and

pamphlets, architectural plans, publicity material, press cuttings and photographs, slides and negatives are found in this collection.

- *Political Collections* contains documents and newspaper clippings on political issues in the 1940s. These include organizations and political parties, labour and trade unions, church unions, the formation of churches and the politics associated with them.

- *Legal Collections* is a collection of legal documents, which includes issues that affected the Civil Rights League founded in 1948. It also has documents about Wynberg student notebooks on legal studies and notices, affidavits, correspondence and many other legal documents.

- *General South African History* consists of papers on a Sociology and Anthropology course on student protest, American newspapers with articles on student protest, and publications dealing with violent and non-violent student protests. It also has the Black Sash Trust Archive and other documents, letters and papers on South African history by different authors (Digitool, 2011).

- *Education Collections* is a collection of documents about education policies and documents in the 1900s. These include the delimitation of functions at university institutions and technical colleges in 1927, the Cape Board of Jewish Education, the Cape of Good Hope Examination Papers, and the 1873 Native Education Commission also known as the Eiselen Commission.

- *Economic Collection* is a collection of documents on trade and industry and economic matters. For example, there are articles from the year 1842 - 1901, which include cashbooks, journals and ledgers of Cape Town firms. There are also books, annual reports, financial records, records of estates during 1904 - 1981, as well as unpublished minutes of evidence of the Customs Tariff Commission of 1934 - 1935 (UCT Libraries, 2011).

- *Cape Town Collection* is a collection of documents and a scrapbook of Cape Town libraries from 1908-1980. It also has the history, constitution and rules, minutes of committee and general meetings, ledgers reflecting the activities and membership of the Cape Town insurance and medical aid society. One may also find taped oral history interviews with people who were squatters in Cape Town in the period 1939-1940. Interviews are in Xhosa and English (UCT Libraries, 2011). There are also transcripts of some of the interviews.

- Under the *Finding Aids,* UCT also has the following collections: Musical collection, Medical Collection, Literary Collection, Legal Collections, Jewish History, Family Papers, Education Collections, Economic Collection, Collection of Women, Cape Town Collection, Botanical Collection, Art and Drama Collections, and Finding Aids. These have not been explained because of the large number of collections noted in Findings Aids and only a few were selected for illustration purposes.

4.3.4 Electronic Theses and dissertations

This is a collection of Masters and PhD student theses and dissertations. The UCT institutional repository at present has a total of 1199 electronic theses and dissertations as it was last updated on the 30th September 2011. The theses and dissertations are from the six different faculties in UCT: Humanities, Science, Heath Sciences, Engineering, Commerce and Law dating from 2008 to 2010. The institutional repository can be found under the UCT library website, under the name DigiTool. The URL is http://www.digitool.uct.ac.za.

4.4 The case study at UCT

The purpose of this section is to discuss the approaches and strategies employed by the researcher to analyse the data collected. Data analysis is the process that involves the use of statistical tools and/or analytical approaches in order to examine the variables under investigation, and finding useful information that assists in addressing the research problem (Bogdan, 1992). He further argues that data analysis brings order, structure and meaning to the

mass of collected data. In this research project the data were in the form of voice notes recorded into the voice recorder, notes in the diary from the participant observations, documentation and physical artefacts analysis. Each interview was recorded and then transcribed. Bogdan (1992) uses the term "field notes" (1992: 107) to refer to the transcripts, notes from interviews and observation. The data analysis of case study methodology is not conducive to statistical analysis (Umit, 2005).

For this research project, data analysis techniques that will be used are the general analytic strategies, case study descriptions and data from the interviews.

4.5 Case Data

In this research, case study descriptions were used to analyse the data collected. The descriptions will first discuss observations that were made at the Digital Initiatives Unit during a month's attachment at the Unit, in the form of **participation observation**. This enabled the researcher to digitise and upload materials on DigiTool (institutional repository) for them to be available on the web. Secondly, the descriptions are based on physical **artefacts** found at the Manuscripts and Archives Department. Their digital surrogates/ images are located at the Digital Initiatives Unit. The images will be counted, checking how many of them are digitized and how many are not. Thirdly, the descriptions also consider **administrative documents** that were found at the Digital Initiatives Unit. Lastly, case study descriptions will be based on the **interviews** that were conducted.

4.5.1 Participation Observation

Participant observation is part of the data collecting techniques as highlighted by Yin (1994). It implies that observers live, as much as possible, with the individuals or organisations they are observing or researching. They try to blend in and take part in their daily activities. During a month of observation at the Digital Initiatives Unit, observations were made through watching, listening and interacting with both the library digitization staff members who are working in the Unit, and with students, who are employed on temporary basis in the Unit. These students help

with the digitisation of materials or projects that have been requested. All the observations were documented. The researcher was actively involved with what is done at the Unit, was given a workstation, and some projects to digitise by the Head of the Digital Initiatives Unit. The typical daily duties at the Unit include the digitisation of images, uploading projects, scanning of theses and dissertations, and scanning on demand.

Scanning on demand refers to scanning that is done by the Digital Initiatives Unit in response to a request by a user of the Special Collections Departments in the Libraries. If a user, who is consulting one of the Special Collections, and finds a document or photograph and requires a digitized copy, the Unit scans it and supplies it to them. There are charges for these scans. The charges are categorised according to who the user is, UCT staff and students, UCT users (in South Africa or internationally) and postal requests. The charges are also based on the type of document the user wants digitised. These range from print to microfilm, digital images, newspapers and magazines, literary manuscripts, books and scholarly journals, television documentaries and feature films, and advertisements for commercial use. (See Appendix D for the attached form on charges).

The Digitization Policy is to scan everything to archival quality. This means that, regardless of what the users' requirements are, scanning is done at the highest resolution and quality, so that the image is later archived and users will no longer have to access the original physical item. If something is requested it is immediately scanned and made available on the libraries' website.

A thesis is digitised by firstly scanning all its pages and later checking if all the pages have been chronologically digitized and that there are no missing pages that have been skipped by mistake. Secondly, by checking if there are any skewed pages or some which have black and dark margins. If there are any, these are straightened using Photoshop and cropped to remove the black and dark margins. Thirdly, the watermark labelled "*University of Cape Town*" is inserted. The watermark is diagonally positioned across the pages at an angle of 45% and the visibility of the words is 10% behind the page. This format is consistent on all the pages of the thesis. Fourthly, the copyright notice page is inserted and the entire thesis document is saved as a PDF. Finally, uploading the thesis on DigiTool by making it OCR text recognisable so that it can be available on the web from the DigiTool institutional repository. The link from the UCT Libraries website is shown by an icon named *Digital Collections*. The URL for this link is

http://www.digitool.uct.ac.za. Other documents are scanned, cropped and saved in the way they appear so that they can reflect their original counterparts. There are no insertions of watermark or copyright notice page to them as is the case with the theses and dissertations.

The researcher uploaded a number of projects on DigiTool. Firstly, a number of theses that are from the Faculty of Law were uploaded. They were first converted from the word document to PDF and their metadata copied into the Microsoft Excel sheet, cf. 2.12.3 for a discussion on metadata. The Microsoft Excel sheet included the following: name of author, faculty and department, year of completion, title of thesis, abstract and system number. These fields are those adopted from the Dublin Core Metadata Schema. In addition to DC the institutional repository project team can also choose to use additional fields based on what works for the institution. DC is the standard used in order for compliance of interoperability with the Open Archive Initiative as discussed in chapter 4 (cf. 2.12.3).

The other projects which were up-loaded on the DigiTool repository by the researcher were "Moving Spirit Images" and "Travelling Light Images" whereby addition of metadata for the projects was done. *Moving Spirit* is a collection of images about Christian African Independent churches in South Africa. The images depict their practices, e.g. how they baptise and celebrate, their way of dressing (i.e. church garments) and their church services. *Travelling light* is a collection of images taken during and after the apartheid era. The images demonstrate the changes in the way of life of the people in Soweto, Johannesburg. The other projects involved digitising images from the TB Davie Memorial Lectures Collection from the previous years to date and thereafter up-loading their metadata.

4.5.2 Physical Artefacts

The Manuscripts & Archives Department has many physical artefacts and some of their images and records are not digitised. The department digitises the images or records based on whether there is a collection that is on demand (cf. 4.5.1.). The physical artefacts that concerned the study are those that will be made available on the web through the institutional repository. These objects include banners, busts, games, and academic dress. The Digitisation Unit is considering digitizing their images by using their Canon 5D camera with a rig and lighting set up.

So far, the Manuscripts & Archives Department have about 147 digitised images and records of physical artefacts. There is one PhD student who is doing her research on physical artefacts at the Manuscripts & Archives Department. When she completes her studies, many of the physical artefacts images and records used will be digitised and added to the collection. At the moment the UCT Institutional Repository has the following artefacts: photographs, manuscripts (i.e. correspondence, minutes of meetings and maps), newspaper clippings, glass plate negatives, lantern slides, slides, negatives, theses and dissertations. These physical artefacts are digitized by means of flatbed scanners and cameras. The Digital Initiative Unit outsources the digitization of film and audio material because they do not have the manpower and equipment to digitize audiovisual materials.

4.5.3 Administrative Documents

The availability of administrative documents was dependent on obtaining access from the Digitization Unit, as some documents are confidential and only meant for the UCT Libraries Senior Management Team. The following administrative documents were found during the attachment at the Digital Initiatives Unit (DIU):

- UCT Libraries ETD Policy,
- Scanning Policy for ETD,
- UCT Libraries Strategic Plan for the years 2010 – 2014,
- Submitting your thesis: a guide for students,
- Metadata fields for still images and theses,
- Proposal Form,
- Copyright page,
- CD cover sheet,
- Delivery Form for materials to the Digital Initiatives Unit (DIU),
- Scale of charges for reprographic and other services for 2011,

- Minutes from the Digital Content Committee,
- Minutes from the UCT Libraries Management Team and
- Inventory of 2011 Visual Archive Collections.

The above documents demonstrated the development and progress of the Repository. The UCT Libraries management has formulated policies to govern the running of the Institutional Repository. They have also adopted international standards that will ensure quality control with regard to the scanning of the images and theses that will be uploaded into the repository:

> "From the start, international guidelines and benchmarks regarding capture parameters and type of metadata were followed. Workshops given by DISA (Digital Imaging South Africa) were helpful in giving an overview of the process and of the practical implications of undertaking a digital imaging project, as were the Northeast Document Conservation Center's (NEDCC) "To film or to scan" workshops. Examining the guidelines of large projects elsewhere, for example the *Colorado Digitization Program scanning guidelines* (now superseded by the *Western States digital imaging best practices*) was particularly helpful" (Dunlop and Hart, 2005:33).

Documents such as minutes of meetings, policies that govern the daily procedures and guidelines on how to execute duties, were closely scrutinised to obtain data, which indicated that the establishment and development of the Institutional Repository is nearly complete. The repository does not yet have an overall institutional repository policy statement document and operates by a policy for submission of print and electronic theses which the Digitization Unit in conjunction with the Library Management formulated. This is a guideline for students on how to submit their theses and dissertations.

Documents help to substantiate the argument that the researcher is putting forward. They often present a researcher with information that may not be available to them at the time of conducting the research and ensure objectivity within the research (Yin, 2003). The major limitation of documents is the reliability and the objectivity of the people who produced them. These limitations were overcome by obtaining views from the people involved in the development and implementation of the Institutional Repository through interviews. The interviews will be discussed in the section that follows.

4.5.4 Interviews

Face-to-face interviews were conducted with those who are actively involved in the development and establishment of the repository and who work at the UCT Libraries in the areas related to information technology, digitization and archiving. Firstly, to orient the study, exploratory interviews were held in March 2011 with the managers or administrative staff of the departmental repositories in UCT (cf. 1.3.1). They were meant to contextualise the digital departmental repositories in UCT. Five people were interviewed who are knowledgeable about their departmental repositories and gave reasons why the departments established the repositories. These included the managers of the following repositories:

a) LawSpace Repository - Faculty of Law,

b) Computer Science Research Document Archive - Department of Computer Science,

c) Manuscripts and Archives Department - UCT Libraries,

d) Open Educational Resources (OER) - Centre for Educational Technology as well as

e) an Honorary Research Assistant whose area of specialisation is Open Access research.

Discussion of these interviews will be incorporated in the interviews conducted between June and July 2011.

Secondly, for more detailed information on the development and establishment of the UCT Institutional Repository, another round of explanatory interviews were conducted between June and July 2011. These interviews were meant to obtain in-depth information from the appropriate sources discussed in the next section below (cf. 4.6). The views and opinions obtained from the interview sessions would be used to validate what is written in documentary sources. This also made it possible to discuss relevant conclusions whilst reflecting the research questions.

Participants were interviewed by means of semi-structured interviews with guided questions are discussed in detail below. Most of the interviews took between 20 – 30 minutes and were recorded for future use. Backup copies of the interviews were transferred and stored on a CD ROM disc. (See Appendix C for the interview guide used during the study).

4.6 Themes from the interviews

From the discussion with the interviewees, a number of themes and topics were deduced. These include responsibility, institutional repository motivation, stakeholders and resources, planning, open access initiatives, consolidation of departmental repositories, strategy for consolidation, institutional repository assessment or success, challenges, and how the challenges were overcome. These and additional issues that arouse during the interview are discussed in detail below.

4.6.1 Responsibility

The interviewees' roles and responsibilities at the unit were established in order to be sure whether those interviewed were sufficiently informed about the development and the establishment of the Institutional Repository. The interviewees included the Head of the Digitization Unit who is in charge of spearheading the Institutional Repository project, and the Manager of Special Collections who works for the Department of Manuscripts and Archives at UCT Libraries and handles the Special and Museum Collections. One interviewee was the Senior Curator of Visual Arts and Film who is responsible for photography and film production collections. He ensures that collections are digitised or if they are already in the digital format they are accessible on the website. Another interviewee was the web designer and technician in the unit, who provides technical support to the institutional repository by uploading digital objects in batch form (thus uploading a lot of digital objects at a time e.g. more than ten to hundred objects). He also solves technical problems in the Institutional Repository and customizes user interface for the website. The Information Technology Specialist who is in charge of maintaining the server and computers was also interviewed.

4.6.2 Motivation for establishing the Institutional Repository

The results from the interviews established the reasons that led to the founding of an institutional repository in the UCT Libraries. These include the need for showcasing special collections as well as electronic theses and dissertations that are for research purposes, and make them accessible on the web. One of the interviewees said that the main reason why the establishment of the institutional repository was set up to make the UCT research output as well as students' theses and dissertations visible. In addition, it was to expose some of the important image

collections at UCT to the world, which are discussed in detail in chapter 4 above (cf. 4.3). When images are frequently handled and not in proper conditions, they deteriorate, and so by digitizing them they are preserved. The goal of the institutional repository is to preserve these valuable heritage collections in the form of images and reports. UCT has some important image collections that attract researchers from all over the world. The repository is a good way of advertising these collections to people who do not know about them.

4.6.3 Stakeholders and Resources

The interviewees highlighted several types of resources that were required by the unit, which include both hardware and software. These are servers, computers, operating system/ platform/software and human resources. One interviewee emphasised that the libraries needed funding for hardware and software (cf.5.2.2). The interviews also indicated that the organizations that played key roles in the founding of the UCT institutional repository were local universities in South Africa such as UP, the University of Johannesburg and Witwatersrand University. These universities shared and exchanged ideas concerning the establishment of institutional repositories at their home institution.

Another organisation that played an important role was Duke University in the USA. Some of the interviewees visited this institution to learn more about its repository as well as other international institutional repositories. They also learnt about digitisation and repository workflows. McGill (2011) describes a repository workflow as …

> "a breakdown of the administrative tasks involved in managing a repository. They allow the various activities involved in the running of the repository to be assigned to the individuals or groups who are best able to deal with them." (McGill, 2011: 79).

There are different types of repository workflows. Examples include workflow to manage user registration, administration, authorisation, permissions, software updates and administrative workflows for maintenance. The most important workflow for repositories is the submission workflow which is frequently used by depositors. This workflow must clearly explain steps to be taken when depositing content into a repository. These steps are gathering necessary data, permissions and files related with the content, and checking the elements or metadata fields required before making the item available on the web (McGill, 2011) At UCT the workflow

process starts with the student submitting the thesis to the faculty administrator. This thesis must be on CD in PDF format. The faculty administrator takes the CD to the manager of the DIU who then hands the CD to the librarian to add metadata and upload it on DigiTool. If the student has not converted the thesis to PDF then the librarian will do so before uploading the thesis.

4.6.4 Planning

The explanatory interviews helped to identify the individuals who were involved in the initial planning of the institutional repository. They were the Executive Director of the UCT Libraries, the Manager of Special Collections and the Head of the Digitisation Unit. The interviews also showed that at the planning stage, no resource audit was done. Secondly, the explanatory interviews established that after the libraries received funding from Carnegie Foundation, the planners were asked to identify what was missing and for them to budget for the resources they needed.

At the planning phase of the Institutional Repository project, software evaluation was one of the major tasks. It took the planners a long time to decide which software to choose. This is discussed in chapter 5 (cf. 5.2.3). One interviewee stated that evaluating which software to use was one of the most daunting tasks they ever faced. It was a long and slow process that involved evaluating both open source and proprietary software. They sought input from different librarians such as the web systems librarian and the librarian who is head of Information Technology. Software evaluation will be explained in detail in chapter 5 section 5.2.3.

4.6.5 Open Access Initiative

Among all the interviewees, only one person addressed the question pertaining to whether the repository will be an open access initiative. The interviewee said that while Digitool is proprietary software, it supports open access and the UCT repository manager has to seek permission from the Ex Libris Company. The other factor that enables open access is its metadata. It uses Dublin Core Metadata which is compliant with OAI-PMH (Open Archives Initiatives-Protocol for Metadata Harvesting), therefore it is OAI-PMH compliant. All institutional repository solutions that use Dublin Core as their metadata standard are compatible with OAI. This aspect is explained in chapter 2 (cf.2.12.2).

4.6.6 Consolidation of Departmental Digital Repositories

The interviewees explained that all the departmental repositories will be brought together under the DigiTool Institutional Repository. According to one of the interviewees, bringing all research work under DigiTool will be an effective way of making UCT research visible. The Law Faculty has been contacted and is in agreement with the proposal. The consolidation of repositories was also confirmed by the manager of the LawSpace repository at the Faculty of Law during an interview in March. He said that the Library management consulted with them about the consolidation of their repository into the UCT Institutional Repository. The reason for establishing a departmental repository was to showcase their research materials online and to promote the open access model of publishing. At the time the library did not have anything in place for such an endeavour.

During the time of the interviews in March, the manager of the LawSpace repository was unsure whether content from their repository, which uses DSpace software, would integrate with contents in UCT libraries Institutional Repository which was to use the proprietary software, DigiTool. However, compatibility had already been established. The Manuscripts and Archives are a department under the Library and would willingly refer the digitisation of their special collections to DIU. The DIU has been holding talks with many departments and has found that there are separate collections all over campus which need to be brought together under DigiTool. At the time of writing, the DIU unit was in the process of contacting the manager of the UCT Computer Science Research Document Archive, about the issue of repository consolidation. The unit has also devised a strategy to use in order to consolidate the smaller departmental repositories.

4.6.7 Institutional Repository Assessment

Although the interviewees gave answers for the questions in this section, they had difficulties answering two questions. One pertained to establishing means of defining the success of an institutional repository and the second concerned the factors that could be used to judge if an institutional repository is successful. One interviewee stated that it required further research. Two of the interviewees said they thought that the number of materials in a repository such as number of theses and dissertations, the number images (visual materials) and how many times

the repository is accessed, could be used as success factors. According to the interviewees, the tracking of hits and downloaded articles can also show the success of the institutional repository. One of the ways of making the repository visible from the UCT Libraries website is to have a bold and attractive icon with the repository's name on it, which would redirect the user to the repository home page when clicked.

When asked what was understood as an indicator of the success of a repository, the Head of Digitisation responded by saying it would probably be determined by putting effective workflows in place. This means that instead of uploading metadata manually, software would be used to upload it in batches, i.e. metadata for fifty to a hundred materials. The other factor would be the time it took for faculties' administrators to take responsibility of uploading theses and dissertations of students in their departments, so that this would no longer be part of the Library Digitization Unit mandate. Lastly, appointing an administrator who would check whether faculties have uploaded the theses and dissertations with the correct metadata, would be another success factor. If all these could be in place and working well, the interviewee was of the opinion that the UCT Institutional Repository would be successful. The interviewee omitted some important success factors such as what users thought about the usefulness of the IR and that higher citation rates could result from theses.

4.6.8 Other issues that emerged

A few of the issues that emerged during the interviews are briefly discussed in this section.

a) The question of whether DigiTool software had a tracking system that could be used after its launch to track user rate received different and contradicting answers. Two interviewees said it was possible while one interviewee said they will explore the possibility but did not really know. The Head of Digitisation stated that if the software does not have the tracking system, then they would have to build their own. Open source software would be used to track the number of downloads. She further stated that such statistics could be used to determine to what extent the institutional repository is used.

b) A question on whether the institutional repository has produced any benefits so far, even though it has not yet been launched. The response was that last year during the open access week a presentation on DigiTool was held which helped to advertise it. Many

students and academics have realised that an institutional repository will be available online and that they will not need to go the library to gain access.

c) Contents from the Computer Science and Law Space repositories make use of ePrints and DSpace respectively. The contents can be integrated into DigiTool without any problem as all the software packages use Dublin Core metadata. The main difference is in the names of the fields. For example, where DSpace and EPrint would use *caption*, Digitool uses the word *title*. The differences have to be accounted for, they cannot be combined or the resources will not be retrievable. Similarly, the metadata fields are not compatible. This can cause confusion to users when they do their searches, because they may not know which term to use in order to find relevant information sources.

d) When asked about positive aspects concerning the institutional repository, the interviewee noted the commitment of the people working in the unit as well as those from other sectors in the library. The UCT Library Management has dedicated some resources, to making staff duties easier and providing a supportive working environment. The Library also provides training to its staff members when necessary. This means that the library tries to train and up skill its staff members. For example, when the Digital Initiatives Unit was established, the Head of Digitisation attended training at Duke University to learn and benchmark best practices on digitisation.

e) Lastly, the researcher asked about the negative aspects concerning the institutional repository. These are listed as challenges in the following section.

4.6.9 Challenges that were encountered

This section will discuss challenges that were encountered in establishing the UCT Institutional Repository.

4.6.9.1 Shortage of staff

All interviewees' viewed staff shortage as the greatest challenge. They stated that they are working extra hard to meet deadlines. One of the interviewees stated that at the moment the

digitisation unit has a digitization librarian, part time staff member and part time students. The staffing needs required by the unit should consist of a digitization librarian, a digital technician; a number of students who can help with digitization, and a metadata librarian whose full time job would consist of uploading metadata.

One interviewee further stated that the shortage of staff is exacerbated by the fact that the unit consists of two departments in one as the institutional repository houses both digital collections as well as electronic theses and dissertations. At UCT the two are combined, yet in other institutions, such as UP, these are two separate divisions. This makes it difficult for the institutional repository to digitize academic papers because they already have backlogs of materials that need to be digitised. It would mean a lot of work and yet they are presently short staffed. There is a way though that could help in digitising the electronic theses and dissertations. *Self-archiving*, by authors as discussed in chapter 2 (cf. 2.9), would assist in reducing backlogs. At UCT it will take time, because the authors from faculties and departments which have never had digital repositories will need to be educated about the repository which is still to be launched. For those departments that already have digital repositories it will be easier as the researchers and authors are familiar with process of depositing their materials on the repository.

4.6.9.2 Shortage of hardware

The shortage of hardware was the second challenge that was discussed by one of the interviewees. He mentioned that the unit does not have enough equipment, as the equipment required is costly. For example, the most expensive digital equipment used for scanning film and video is the Hasselblad flextight 5 Scanner which costs about R200, 000.00. It can scan all formats from 35mm films to 4 x 5 inch negatives. Other small scanners for digitizing photos, negatives and paper include flatbed scanners. These are also expensive but cost less than the film and video scanner. Another interviewee added that the unit's equipment is constantly used, averaging at eight hours per day. If a scanner goes out to be serviced, the production level drops as no work is done during the time. This problem also contributes to the backlog of work and it was proposed that equipment should be hired when others are being serviced.

4.6.9.3 Lack of flexibility of DigiTool

The third challenge raised was that the DigiTool software, as a commercial product, does not allow changes. The interviewee explained that proprietary software is built and compiled on programme codes so the company has to be called for the codes to be customised and adjusted. For example, if the interviewee wants to simplify the uploading procedure, he cannot do that himself and would have to call the company. Similarly, if a feedback column in their website is required where users are given the opportunity to comment like in a blog, they have to call the company for the programme codes. The interviewees expressed that this is a limiting factor. Another interviewee added that an obstacle they face with the DigiTool software concerns the arrangement of collections. This means that whenever they add or upload a collection into DigiTool, it makes it a separate collection, instead of merging it with the already existing collection of the same name or just adding the newly uploaded contents to the existing one. The interviewees would prefer a situation where added collections are merged with the already existing ones in DigiTool, thereby not having many separate entities of the same. For example, at the moment they have different thesis collections for Law, Science, Humanities, and Commerce instead of just having one collection on theses. During the time of the interview the Head of Digitization complained that she had once logged a call with DigiTool personnel and that it had taken three weeks for them to get a response.

4.6.9.4 The problem of archiving audio visual materials

Two of the interviewees mentioned that another challenge they foresee is how the institutional repository will archive audiovisual materials to be watched at a later stage. This is known as ***streaming*** which refers to the playing of sound or video over the Internet or a computer network in real time. The problem was how to decide which ones would be shown on the website and whether they would show a whole video or short clips of it. Another problem with the audiovisual materials was the issue of bandwidth. The Digital Initiatives Unit has many films and videos, but South Africa does not have the bandwidth that allows for the streaming of online videos such as those seen on ***YouTube***. In addition, bandwidth is very expensive for African countries.

4.6.9.5 Insufficient training

The fifth obstacle concerned training issues. The interviewees stated that Digitool provided basic introductory training which involved uploading the digital objects, online tutorials and training on how to manage an institutional repository. One of the interviewees stated that this basic training is insufficient and did not equip him with all the technical training he needs. He had to learn many of the new technologies on his own as he was used to working with open source software. The interviewee stated that he is not a librarian and that he is learning how to use the MARC formats, Dublin core, MEDS and MODS standards. He is not familiar with these standards and wishes to be trained on the practical aspect of dealing with Digitool software. He also had to learn how to extract metadata from the digital objects that were uploaded on Digitool. This included extracting metadata from photographs and videos while learning how to use the software. He was of the opinion that open source software would have been the better option as he is more familiar with it than proprietary software. The UCT libraries sent some staff members for training in Duke University but it was limited to supervisors.

In a discussion in chapter 5, section, 5.2.3, further explanation will be given why UCT libraries management chose DigiTool over open source software. One of the reasons was that it was based on the premise that DigiTool offered ongoing technical and training support in the form of a contract with Ex Libris. The arguments were that the UCT Libraries did not have enough technical expertise or sufficient staff members with the right qualifications and experience on campus to deal with open source software, trouble shooting and customization. Ex Libris, the company which offers Digitool, undertook to provide these services. Whenever they faced a problem with DigiTool, they could log a call for help and it would be immediately attended to. In theory, this seems to have been the biggest convincing factor for the UCT Libraries Management. In practice, however, this seems not to be the case. The people who are actively involved and working with the software on a daily basis complain about a lack of training, the lack of flexibility and the length of time it takes for Ex Libris personnel to provide feedback on reported problems.

4.6.9.6 Office Space

The sixth challenge was that of ***limited space.*** The interviewees complained that the Digital Initiatives Unit is situated in a very small area. It would be ideal if moving to a bigger location that can accommodate all their machines and equipment were possible. The Digitisation Unit occupies a space of about 20 square meters. The researcher also observed the issue concerning space during the month of participation observation at the unit. There were times when the researcher was asked not to come to work because there was digitization to be done on demand that the work space was needed.

4.6.9.7 Other issues

Although it is not part of this research, the final challenge noted is that of the *competition for African heritage*. One of the interviewees expressed a concern that the world is interested in African heritage especially South African heritage. First world universities have the financial resources to buy the collections on African heritage, which the African universities cannot afford because of the lack of funding. Therefore, this poses a threat of losing the African heritage to first world institutions. The lack of a national policy on digitization in South Africa makes the challenge of competition for African heritage difficult to manage.

How the above challenges were approached are discussed in the following section.

4.7 Resolution of challenges

Although the Unit faces these challenges, it is working to ensure that the research materials, images as well as theses and dissertations are digitised and placed online for users. The resolutions to these challenges are described as follows:

- ***Shortage of staff*** will be solved by appointing a full time member of staff who has been redeployed from another section in the library. This person will do the digitization and adding of basic metadata. The plan is to utilize the member for scanning on demand, uploading theses and dissertations, and the backlog of these activities. There is a huge backlog of materials that need to be digitized.

UCT Libraries Management plan to assign one staff member to the Unit. This person will only be based there for three hours per day to help. He is from one of the sections in the library and will attend to urgent backlogs.

One way of curbing the problem of staff shortage at the Digitisation Unit would be to introduce ***self-archiving***. This is beneficial to both the scholars/researchers, the institution and society as a whole. What the UCT repository manager and her team need to do is educate UCT scholars about self-archiving and make them understand its benefits. Secondly, they should provide support services to these scholars in order to encourage higher rates or self-archiving participation (Kim, 2010). Support services could include having a clerical assistant whose job would focus on archiving for the scholars, as many of the scholars find it time consuming or do not have the technical skill to do it. If done properly they may participate in depositing their research materials on the repository.

- ***Shortage of hardware*** – Library Management is still working on how to deal with the challenge of equipment shortage. The Head of Digitisation has proposed that as the equipment is expensive, an alternative option may be to hire equipment when the unit machines has are out for service. There were times when the researcher would be asked not to come to work, because there was digitization to be done on demand. The part time students who would be doing the work had to use the workstation or computer and scanner, as there are only two of them.

- ***Insufficient training*** – One of the interviewees mentioned that he relies on his supervisor, the Head of the Institutional Repository for help whenever he counters a problem.

- ***Limited space*** – the Head of Digitisation has already requested the Library Management to look into the issue and see if it is possible for them to be moved to a larger location.

- ***Another challenge*** includes the lack of bandwidth for streaming videos and films. This challenge is beyond the control of the Digital Unit, the UCT Libraries and even the University. This is a national issue that can be dealt with at Cabinet level or by the Department of Science and Technology or the Department of Communication. The University and the Libraries would have to do the best they can although faced with such a challenge.

4.8 Conclusion

This chapter presented a case study of the establishment of the UCT Institutional Repository by firstly outlining the background of the project. The project was developed into a repository after some funding from the Carnegie Corporation of New York. It further outlined how the repository started, how its contents and the study data were collected. This was done through a number of techniques including participation observation, physical artefacts, documents and interviews.

These techniques provided detailed results of the data that was collected from interviews and also confirmed the findings from the participation observation, physical artefacts, and documents. The interviews strengthened the findings that the development and establishment of the institutional repository was in progress. The interview discussion was based on the themes and topics that were deduced from the case study. Resources had been gathered, structures put in place, planning done and strategies devised, members of staff redeployed and digitization started. It concluded reviewing the challenges encountered and how these were overcome.

The next chapter will compare the establishment of the UCT institutional repository with that at UP. The factors that could be followed when establishing an institutional repository will be considered as well as how each institution dealt with them. It will also discuss the criteria used to evaluate the success of a repository.

CHAPTER 5

CRITERIA FOR A SUCCESSFUL INSTITUTIONAL REPOSITORY: A COMPARISON BETWEEN UCT AND UP

5.1 Introduction

In this chapter, the discussion will focus on a comparison between the institutional repositories at UCT and UP. The research on UP will be based on published work, as time could not permit the researcher to travel to Pretoria for some interviews. The comparison will highlight the factors to consider when establishing an institutional repository. The discussion will include the criteria necessary for a successful institutional repository. This chapter is partly based on, and expands in detail, the issues discussed in a paper co-authored by the researcher, Annah Macha and her supervisor, Associate Professor Karin De Jager. The paper, entitled "Two South African Institutional Repositories: a comparative overview," was presented at the *14th International Symposium on Electronic Theses and Dissertations,* held on the 13-17 September 2011, Cape Town, South Africa and published in the Proceedings. Furthermore, this chapter concludes by providing a way forward for UCT and discusses UP as a benchmark for other universities with regard to the establishment of institutional repositories in South Africa and the Southern Africa.

5.1.1 Background of the institutional repository at UP

In order to better understand the UP's institutional repository and know the background information of the university, a personal interview was held with the manager of the repository during the *14th International Symposium on Electronic Theses and Dissertation* conference. The manager of the University of Pretoria electronic Theses and Dissertations (UPeTD) explained that UP is one of the largest universities in South Africa and is one of the leading research universities in both South Africa and in Africa. The university has six campuses, with nine faculties, and a business school. According to the latest statistics from the Bureau for Institutional Research and Planning office, the total student population at UP in 2010 was 63 500 of which 1456 were postgraduate students (Olivier & Louw, 2011).

5.1.2 Establishment of an institutional repository at UP

As a pilot project, UP started making both theses and dissertations available online during 2000 (UPeTD website). This was the result of a visit by Prof Ed Fox of Virginia Tech USA in 2000, who demonstrated the e-thesis specific ETD-db software to the library management of UP. Virginia Tech originally produced the ETD-db software. Shortly afterwards, the Library was tasked with establishing an institutional repository for which the open access ETD-db software was used (Hammes, 2010). At the end of 2002, the repository contained 39 theses and 26 dissertations. In 2003, a policy was adopted by Senate to make the submission of electronic theses and dissertations compulsory. For this, a new server was bought. In 2004, based on its success, UPeTD expanded in order to create a repository, especially for research collection. The management of the UP institutional repository soon discovered that the institutional repository required full time clerical staff.

Another milestone at the University of Pretoria was the establishment of a second repository, which is the institutional repository of the University, UPSpace. UPSpace is a repository which was set up mainly for the management and dissemination of digital research materials donated to or created by the research community for their special collections. It has three main collections, which are popular research, scholarly materials, and historical/archival materials. The project of setting the UPSpace repository started in 2004 and was finished and launched in 2006 (Bothma et al., 2008). The UP project team met and drafted a proposal to the university management outlining the approach they wanted to use to start the project. This was the project plan, costs, project phases (analysis, design, development, implementation and evaluation) and the outline of institutional repository. In 2005, the team went further to evaluate the different open source software packages for the proposed institutional repository and finally chose DSpace. By 2006, it became operational. According to Pienaar and Van Deventer (2008) UP also has OpenUP, which is a sub-collection of the larger UPSpace collection. This "repository collection offers open access to the full text of research articles published by staff, students and affiliates of the University of Pretoria" (Hammes, 2010: 2). OpenUP houses e-prints as well as peer reviewed and published research articles (Hammes, 2010).

UP therefore has two repositories, UPeTD and UPSpace while OpenUP is part of UPSpace (Hammes, 2010). UPSpace is a collection that supports the Green route while OpenUP offers authors two routes of publishing their research output. Open Access self-archiving, also known as the *Green route,* allows authors to submit their post-prints to the repository. Open Access publishing known as the *Gold route* allows authors to publish in open access journals (Hammes, 2010).

During an interview, the UPeTD manager explained that the repository offers four levels of access to the UP electronic theses and dissertations. The option chosen is mainly determined by the supervisor and the student, who is the writer of the thesis. However, copyright belongs to the University. The four possible levels are:

- "Worldwide availability;
- Limited access within the UP domain;
- Restricted access for a period of two years – usually when the student and the supervisor first want to publish from the thesis. The two years give them time to complete the publication process of the research articles, after which it will then be open to everyone; and lastly,
- Total embargo to prohibit access to confidential or sensitive information" (Olivier and Louw, 2011: 3-4). This means that the contents of thesis are not seen at all, the metadata and the abstract can be made available, but the publication is banned for security or copyright reasons (Olivier and Louw, 2011). According to the UPeTD there are very few theses that undergo this access option. Examples of theses and dissertations in this category are mainly from the political science or medical science fields where students have sensitive political issues or photos of patients in the theses (Olivier and Louw, 2011: 3-4).

At UP, the faculty postgraduate officers are charged with the responsibility of collecting the students' theses and dissertations in both hard copy and in electronic format. They then deliver the electronic copies with submission forms to the Open Scholarship Office. The postgraduate

administrators are also charged with the responsibility of taking hard copies to the library after graduation for shelving (Olivier & Louw, 2011).

In 2005, UP became a member of the Networked Library of Theses and Dissertations (NDLTD). This international body provides access to the entire world's theses and dissertations and holds annual conferences. It also rewards electronic Theses and Dissertations (ETDs) that use technology in innovative ways. UP has contributed to nine workshops in South Africa, Botswana, Lesotho, Ghana and Ethiopia, sharing expertise and enthusiasm and helping colleagues to start their own operations (Hammes, 2010). They assisted the Council of Scientific and Industrial Research (CSIR) to develop a digital repository (Pienaar and Van Deventer, 2007). So far, staff and interested parties outside UP have given seven presentations at different ETD conferences and workshops about UPeTD. UP is actively involved in open scholarship, library quality, and facilitates open access of information for research with its two repositories. Furthermore, it indicates that to a considerable extent, UP is the leading university in South Africa in this field (Hammes, 2010). It has also become known that UPeTD will soon disappear. As a separate entity as there are developments under way to transfer and integrate the contents into UPSpace (Pienaar, 2012).

5.2 A comparative analysis of the UCT and UP institutional repositories

In the next section, the institutional repositories of UCT and UP will be compared. UP was chosen for the comparative analysis based on its well-established repository and success story. UP has identified a number of factors listed below which may be taken into consideration when setting up an institutional repository (Smith & Pienaar, 2007). The following six factors are used to compare the two institutional repositories and can be referred to as prerequisites for establishing an institutional repository.

a) Identify role players,

b) Address the issue of resources,

c) Evaluate software,

d) Formulate a policy for the Institutional Repositories,

e) Restructuring the library to accommodate change and

f) Licensing

These six factors have been discussed in depth in the section 2.9 and here it is just a comparison of how UP and UCT have been able to meet with the requirements.

5.2.1 Identify important role players

With regard to identifying role players at UP was able to identify and included "the Head of the Department of Information Science, subject librarians who were comfortable with new innovative tools, a metadata specialist, a digitization specialist and Information Technology [IT] staff" (Smith & Pienaar, 2007:4).

At UCT, it was shown that the main role player is the Head of the Digitization Unit who spearheads the UCT Institutional Repository project. Her role is that of project manager and therefore is the Institutional Repository champion. She is supported by various staff members, such as subject librarians, metadata specialists, and IT specialists.

5.2.2 Address issues of resources

At UP, the project of establishing an IR was carried out by firstly finding out which resources were needed. This was done through conducting a survey, which helped identify the resources the team would need, and those that UP had. The survey revealed that UP had a server, funds and supportive staff members especially from the library management, specifically the Assistant Director of eInformation Strategy & Knowledge Management.

Unlike UP, UCT Libraries started with a small trial project to digitize some photographic collection they had at the time in 2001. The project was not intended to establish an institutional repository; it was mainly to digitize photographs. It was only in 2009, when UCT got funding from the Carnegie Foundation, that when the library decided to expand the digitization project to

include digitization of other documents like theses and dissertations and turn the unit into an institutional repository. Funds from the Carnegie Foundation were meant to support research in the library, hence establishing an IR was a suitable project.

5.2.3 Evaluate software

At UP, open source software and open standards based on the UP Information Technology framework. The open source software, ETD-db was chosen for the theses and dissertation repository (Pienaar and Van Deventer, 2008). In 2004, the UP digital repository project team evaluated software in order for them to establish a second repository, UPSpace. It took the project team over six months to complete the evaluation of several open source and commercial software platforms such as Greenstone, Innovative, Fedora, ePrints, DSpace, I-Tor, MyCoRe, CDSware, ARNO, Innopac, Mellennium and WebCT. DSpace was eventually chosen for the digital repository, because it fitted the UP Information Technology architecture and supported its workflow (Bothma, Pienaar and Hammes, 2008). The other reason for choice of DSpace was based on the software evaluation documentation entitled "*Motivation for the deployment of DSpace, an Institutional Repository System, at the University of Pretoria*". This document can be found at url; www.dspace.up.ac.za/motivation.doc.

In 2009, after receiving the grant from the Carnegie Foundation, the UCT library team, which was in charge of establishing an institutional repository started evaluating both open source and proprietary software. The evaluation took into consideration the software packages that were mostly used by other universities in South Africa and those that were being used in countries such as America, United Kingdom, Canada and Australia. In South Africa, it was found that many universities were using DSpace and ePrints, which are both open source software, while the proprietary software used, were Luna Imaging and ContentDM.

The establishment of the UCT institutional repository was spearheaded by the UCT Digital Initiatives Unit, which assessed five software packages. These packages are DSpace, ePrints, ContentDM, DigiTool and Omeka. The fifth software package Omeka was assessed and evaluated but was later eliminated as it was found not to be widely used in many countries. The assessment and evaluation selected DSpace and ePrints as, open source software and DigiTool

and ContentDM proprietary software for further assessment. Neither of the two open source packages were chosen, since UCT was interested in a software that could best show photographic materials online. The team then settled for proprietary software DigiTool, though DSpace could have been the open source software of choice.

The main difference is that DSpace does not show a visual overview of the photographic collections. This is only shown by proprietary software. For example, when one hovers the cursor over an image in DigiTool, one would immediately see the information about the image, in the form of a visual overview (visual display jpg). This will for example, show *Type* - JPEG, *Dimension* - 960x 720 pixels and *Size* - 74.3 KB. For open source software one has to click on several links before being able to view one image. It was also found that it was possible to customize the look and feel of the interface on both packages, although there was yet no instance where this had worked effectively. Another reason why DigiTool proprietary software was chosen over open source software, was that the UCT library was of the opinion that it did not have the capacity to develop sufficient local expertise to sustain and maintain the installation of DSpace and EPrints.

ContentDM and DigiTool were similar in functionality and price. DigiTool was finally selected, as it is a reputable, vendor-supported tool, adopted by many major universities. UCT library had a long-standing relationship with Ex Libris (the company that created DigiTool). DigiTool would integrate well with the UCT online catalogue, Aleph, and UCT portal, PRIMO by Ex Libris, the new UCT search interface. Since UCT digital Initiatives Unit started using DigiTool, the contents in the repository began to increase. In addition to the reasons above, DigiTool offered UCT ongoing technical support in the form of a contract with Ex Libris.

It is important to note that DigiTool can support open access as it is an Open Archives Initiatives-Protocol for Metadata Harvesting (OAI - PMH) compliant enterprise solution for academic libraries or research institutions (Stevenson and Hodges, 2008). DigiTool was designed for creating, managing, preserving and sharing digital collections. It also supports Dublin core. The Ex Libris website states that "the open-access platform also facilitates the sharing of scholarly information in a library's user community by integrating digital collections with

institutional portals, e-learning systems, and other library systems" (Ex Libris website, 2011). Although it is proprietary software, it meets the needs and requirements of open access. The UCT institutional repository will be an open access initiative in spite of being based on proprietary software.

5.2.4 Formulate Policy for the Institutional Repositories

At UP, the ETD repository is governed by a policy which has been approved by the University Senate, and is managed by a team consisting of the University Librarian, the Systems Librarian and the Institutional Repository Librarian (Smith & Pienaar, 2007). The policy was approved in 2003 and, according to Bothma, Pienaar, and Hammes, was compulsory: "all students registered for a Masters or a doctorate had to submit their research in electronic format (in both MS word and PDF format), either directly to the repository or on a Compact Disc to the Faculty Administration, before graduation" (2008: 276). Students were provided with guidelines and tutorials on how to submit the theses, which can be found on the UPeTD website http://upetd.up.ac.za/authors/publish/index.htm. According to the UPeTD manager, all students who self submit their theses, are rewarded by having their theses processed first. There is also a policy for archiving of journal articles and other research output as well as a policy for other collections as well. As for UPSpace the policy that stipulates mandatory submission for research articles was since adopted in 2009 (Olivier, 2011)

UCT has a policy for the submission of print and electronic theses only. According to the General Rules and Policies Handbook of the University, every postgraduate student is required to submit one final, corrected copy of the thesis in a printed format and one final, corrected copy in electronic format. The electronic copy will be accessible from the Libraries' website as a PDF document through the DigiTool website. This is via a link in the online library catalogue (Aleph) while the print copy (bound and catalogued) will be shelved in the main library. The other catalogued copy will be housed in the African Studies Library and will serve as the archival preservation copy. It will not be available for loan (DigiTool website). The policy also has guidelines for students on how to submit their theses and dissertations, but does not yet have an overall institutional repository policy statement.

5.2.5 Restructure the library to accommodate change

As mentioned above the University of Pretoria has two repositories and for both of them the library had to restructure and accommodate change. To remove any possible uncertainty amongst role players, each role was clearly defined and communicated. Certain responsibilities were assigned to these roles so that all staff members were familiar with exactly what was expected of them (Smith & Pienaar, 2007). Involvement in repository activities is part of the official job description of cataloguers and subject librarians. The librarians' performance is monitored on the performance management system.

Firstly, with the UPeTD, staff roles and responsibilities changed as theses and dissertations had to be digitized. The library needed more staff who could do the digitization. In 2007, the library increased the number of subject librarians who were to be involved in digitization to 18 and increased metadata editors to 12. Nine of the library leaders were elevated to become collection managers (Smith & Pienaar, 2007).

Secondly, when UPSpace was established, it changed the way day-to-day work at the University of Pretoria library was performed. Bothma et al (2008) noted that the librarians had to undergo training in order to cope with the new job requirements. Some librarians decided to enroll themselves for courses to equip themselves with relevant skills. Those who were trained by the library were trained for the following areas: collection, open access managers, digitization, and metadata specialists. The change that was brought by the repository also brought some challenges such as, the staff had to:

- "support the strategies and objectives of the institutional repository projects;
- influence the mindsets of colleagues in the library and also of researchers that may be hesitant to accept these new ideas; encourage knowledge transfer,
- motivate others to share, learn, apply their new knowledge;
- communicate new ideas and new ways of working effectively to their colleagues;
- identify new opportunities; and
- mentor others that may be struggling in the new work environment" (Smith and Pienaar 2007).

Although the library faced the above challenges, communities of practice emerged and the sharing of knowledge started to take place, for example, the UPSpace project team started to collaborate with information specialist from the CSIR (Bothma et al., 2008).

UCT Library management is in the process of restructuring roles and responsibilities of staff in order to accommodate the change that will be brought about by the establishment of the Institutional Repository. Existing staff members will be offered new roles and new staff will be hired for the new roles (Macha and De Jager, 2011: 6).

5.2.6 Licensing

As noted in chapter 2.9.6. "in order to ensure open access to its collection, an institutional repository needs to be registered with open access harvesters such as OAISTER, Registry of Open Access Repositories (ROAR), Open DOAR and Google Scholar. This allows the content to be harvested by these search engines for visibility. These search engines would then be able to pick contents from the institutional repository (Stevenson & Hodges, 2007). Both repositories at UP (UPeTD and UPSPace) registered with ROAR, openDOAR, Google, and as a publisher at Google Scholar, as well as with the DSpace community.

After setting up the institutional repository, the project manager at UCT would also need to register the institutional repository with open access harvesters such as those noted above. After registering, the OAI service will then perform conformance testing to ensure that the UCT institutional repository complies with OAI-PMH. If validation is successful, the repository will be added into the registry. The OAI will also periodically test the repository for conformance. If the analysis fails, the repository will be removed, and a notification email is sent to the administrator detailing the reason for removal. This ensures the integrity of the OAI registry and the repository interface (McGill, 2011).

5.3 How UCT IR complies with the criteria for a successful institutional repository

The criteria for the success of an institutional repository would normally be taken from its statement of purpose. An institutional repository can only be successful if it fulfills its stated purpose which includes "preserving information and making it available for a designated community" (Thibodeau, 2007: 1).

5.3.1 Content

The success of an institutional repository is seen through the number of contents it has. Below are the figures that show how much content is there in the UCT institutional repository and the numbers in the UP repositories both from the UPeTD repository and from the research repository, thus the UPSpace.

According to Olivier and Louw (2011), UPeTD had a total of 6621 materials when it was last updated in August 2011. UPSpace which, is the second repository at the University of Pretoria houses research articles, conference papers, datasets, theses/dissertations and multimedia materials has a total number of 13864 items as up dated on the 24th of August 2011 from the (OpenDOAR, 2012). UCT currently has 4063 materials, which was last updated on the 30 September 2011 (DigiTool website).

5.3.2 Use

In 2009, UP Electronic Theses and Dissertations had 4317706 files downloaded and received messages from users testifying to the usefulness of the repository system (UPeTD website). A survey done by Hammes (2010) indicated that its alumni PhD students, who were experiencing rewards in their careers because the ETD had made their theses visible, appreciate the ETD. According to Olivier and Louw (2011) in August 2011 the most read item from their repository was accessed 30 384 times. UPSpace repository has on its home page a section where one can a view usage statistics. This section is divided into three categories. The first category is the *general overview* of how the repository is used which shows statistics of items archived, bit stream views, items views, collection views, user logins, search performed, and OAI requests.

The second category show the *statistics of archive information* which shows the content type viewed, was the user viewing a thesis or a research paper, the number of items viewed. Lastly, the *number of visits per article*. UPSpace tracks all the daily, monthly and article usage statistics, and after that archives them.

During an interview, UCT's LawSpace repository manager stated that the Faculty of Law receives emails from readers outside the country, who use and appreciate the LawSpace repository, asking why recent student theses have not yet been uploaded. According to the Head of the Digitization Unit, the UCT institutional repository intends to track the use of Institutional Repository through daily downloads and hits.

5.3.3 Submissions

At UP, a policy was adopted by Senate in 2003 to make submission of theses and dissertation compulsory, thus in the UPeTD repository. The UP Senate voted for a mandate for deposits indicating that there is top-level support for the institutional repository establishment (Gray, 2009). At UP according to Pienaar (2012) most theses, dissertations and journal articles are uploaded from a born digital format. Workflows for different staff including assigning metadata and quality control are important. Students upload some of these documents while the academic staff self-archive their research publications. UPSpace also has a mandatory submission policy endorsed in 2009 for research articles. Software that has been specifically developed for these workflows exist. DSpace is an essential tool in this context. DSpace can also handle multi-media materials to some extent.

At UCT, the Library policy deals with the submission of electronic theses and dissertations and digital collections. It is compulsory for students to submit in a digital copy of their thesis in the form of a CD ROM so that they could be uploaded into DigiTool. The policy regarding submission and deposits into the UCT institutional repository still has to be formulated, approved and then passed on to the UCT Senate for final approval.

5.3.4 Advocacy

The way the UP advertised the electronic theses and dissertation repository to the university community was first by forming a campus committee. The mandate for this committee included creating a framework within the university policy, making guidelines and resources accessible to students, teaching staff and administrators. The committee also created a website, produced brochures, and organized regular anti-plagiarism campaigns held at different spots on campus (UPeTD website).

The UPSpace project team came up with a number of marketing strategies, which they put in place in order to alert the UP community of the existence of their repository. These are the strategies they implored: holding meetings, forming a UP library community, having presentations to faculties and departments, producing a UPSpace online newsletter, and listserv. They also produced a programme for orientating new students and staff, started an online brochure and started to give and show statistics to faculties of members who used the UPSpace repository. They also used the UP Department of Marketing to help them with the marketing of the repository (Smith, 2008). In 2006, in the month of July UPSpace Project team held presentations to different faculties and departments in the university, *"UPSpace: an institutional research repository for the University of Pretoria"* which was meant to teach academic staff, researchers and students of their IR and its benefits (Smith, 2008).

UCT has not yet reached this stage. The institutional repository is still in its infancy and it is suggested that marketing plans will be put in place and executed in preparation for the launch.

5.3.5 Support

The Senate in UP voted to mandate deposits and in so doing indicated top-level support for the establishment of the institutional repository (Gray, 2009). This is a sign of constituent support that UPeTD receives from the University Senate. This means that the institutional repository largely relies on mandates. This was supported by Harnad and McGovern (2009: 28) who stated that "mandates provide the only means for achieving a successful institutional repository," as UP also demonstrated. UPSpace is also supported by the University management team. UPSpace

took a step further to support its users not only to get support of the university community. It established a landline help desk where all users needing help could call for free eight hours a day and in the five working day. The online help service was 24 hours every day, it entailed Viewlets such as Macro media Captive, Power Point presentations and open Office. There was also an online tutorial, a user policy guide and pamphlets and email address to send all request and questions at dspace@up.ac.za (Smith, 2006).

At UCT in 2009, the library obtained funding from the Carnegie Corporation to support research initiatives. This funding helped UCT to start setting up the institutional repository and to purchase the DigiTool software. It is not clear how much constituent support it will receive. What is clear is that students are mandated to submit an electronic copy of their thesis to their faculty administrators before graduating.

5.3.6 Influence

Influence refers to providing assistance to other institutions in establishing their own institutional repositories. This is similar to what Thibodeau (2006) calls collaboration. It is about how the institutional repository associates with others in the country, region and in the world. Institutional repositories are encouraged to work in collaboration with others to demonstrate that they are successful. If an institutional repository works in isolation, it might signify failure because it does not share and transfer knowledge. A more collaborative arrangement would exist where separate institutions decide to work together to leverage each other's strengths and where they form or join consortia for such purposes (Thibodeau, 2006).

This is already happening at UP, which is actively involved in helping other institutions set up their institutional repositories. It assisted CSIR as well as the Southern African Regional Universities Association (SARUA) representatives who visited them in 2007. Although the UCT institutional repository is in its infancy, it is already active in supporting other institutions who wish to learn about the theory and practice of digitization. The Head of the Digitization Unit has delivered workshops to librarians and archivists. The Kimberley Digitization Workshop and the Digitization Workshop for Library of Parliament librarians in 2010 are examples of this. At UCT, she teaches digitization to Post Graduate Diploma and Honours students in Library and

Information Studies. The UCT institutional repository staff has also advised various campus departments on metadata standards and the implementation of good metadata practice.

5.3.7 Interoperability

UPeTD is registered and therefore interoperable. It registered with OAISTER, Google Scholar and DSpace. UPSpace is registered with a number of international harvesters and registries such as OAISTER, Institutional Archives Registry, DSpace Instances, DSpace wiki and Open Archives Institute, Google, and Google Scholar, so that its contents can be harvested by them and shared worldwide. Although UCT has not yet reached this stage, it will register in due course and become interoperable.

5.4 A critical summary: UP as a benchmark and the way forward for UCT

From the comparison between institutional repositories at UP and UCT it is clear that although both institutions have repositories, they are not similar. Firstly, one can look at the differences between how UP and UCT's repositories started. UP's first repository focused on housing electronic theses and dissertations (ETD) and later, another repository, UPSpace, was for research output and for the institutions' academic community. Open UP is a sub collection of UPSpace.

UCT's repository does not primarily focus on electronic theses and dissertations like UPSpace. It is for all institutional digital objects, including research output such as, research papers, data sets, images, digital collections, as well as electronic theses and dissertations. At the moment, the UCT institutional repository focuses on special collections and heritage collections found in South Africa. It has one of its image collection listed under UNESCO'S Memory of the World register as a documentary heritage of international importance. The UCT Libraries strategic plan for the years 2010-2014 also shows that the repository's main focus is on:

> "strengthening the Special Collections and Digitization Units to attract additional heritage collections and to enhance capacity to digitize special collections and research materials. In

particular, focus on providing adequate and appropriate space and added curatorial expertise in specialized materials, such as film, video, photographs, and music." (Rapp, 2009: 6)

Secondly, one can also look at the different types of software that each repository chose. UP chose open source while UCT chose proprietary software. This was attributed to viewing its contents which will be dominated by images. Software that displays images clearly and satisfactorily was necessary. Proprietary software provided this, while open source software could not. Photographs and other visual material are not displayed in a satisfactory manner in open source software such as DSpace or EPrints. DSpace does not have the same visual impact as the products DigiTool or ContentDM. The main difference is that with open source software one does not get a visual overview of the photographic collections as one would get with proprietary software.

Thirdly, UP has a mandatory policy on submission for both students and staff. This has been in operation for some time now since 2003. It helps its repository to have a large volume of content and in 2010, the repository was ranked number 95 in the world as stated by Hammes (2009),

> "UPeTD recently celebrated its first decade with its best rating in the January 2010 *Ranking of World Repositories:* out of 400 best institutional repositories worldwide it is no. 95 in terms of content, size and visibility as taken from http://repositories.webometrics.info/index.html" (Hammes, 2009:1).

This is not the case in UCT, as it has a policy for students' theses and dissertation only.

Lastly, UP can be taken as a benchmark for other institutions including UCT because of the success of its electronic theses and dissertations repository. It is among the early institutions in South Africa to set up a repository and has assisted other institutions in establishing their own. It sets a good example when it came to establishing institutional repositories in the country, and holds workshops and seminars to influence other research institutions to follow in the movement of open access and scholarly publishing. It assisted CSIR and expressed interest in working with and helping the Southern African Regional Universities Association (SARUA) representatives

who visited them in 2007. It is not only a role model to South African institutions, but in the region as well.

5.5 Conclusion

This chapter presented a comparative overview of the establishment of the UCT institutional repository and compared it with that of UP, highlighting the factors to consider when establishing an institutional repository. It was shown that UCT started its repository after UP, and its advantage is that it is now able to benchmark from the most well established institutional repository in South Africa. The importance of identifying role players, addressing issues of resources, evaluating software, policies for the institutional repository, restructuring the library to accommodate change and licensing were considered. The UCT Institutional Repository was then compared with that of the UP.

The chapter discussed the criteria for a successful institutional repository, which could be used by various universities if they wish to do so. Criteria for a successful institutional repository have also been discussed in order to show what needs to be done in order for an institutional repository to be successful and be of value to its stakeholders. It has also been shown that institutional repositories can play an important part in a university, as they collect and store all the university's research output and in so doing, assists with the preservation and archiving thereof. Institutional repositories are also important in enhancing the visibility of the institution in Africa and the world. In summary, the evaluation and assessment of institutional repositories is currently not standardized.

In the next chapter, the findings are discussed according to the criteria for the success of an institutional repository, a conclusion of the whole thesis is reached, recommendations made, and finally, topics for further research are proposed.

CHAPTER 6

DISCUSSIONS OF THE KEY FINDINGS, CONLUSION AND RECOMMENDATIONS

6.1 Introduction

This study set out to investigate the development and implementation of an Institutional Repository at the University of Cape Town. The study shows that the UCT Institutional Repository would address the demands of digital preservation, curation and creation and transfer knowledge. It also has the potential to make UCT visible to the world as a leading African teaching and research university.

This chapter will discuss and interpret the study's key findings from the data analysed in chapter five. The chapter will also summarise findings and address research the questions.

6.2 Discussion of key findings

The findings will be discussed according to the criteria for success of an institutional repository. The study identified factors that could be used to assess the success of an institutional repository. It is suggested that if these factors are to be taken into consideration in the UCT Digitization Unit, their repository could be regarded as successful. It was shown that the development and establishment of the Institutional Repository at the UCT is nearly complete. The factors that seem to make a repository successful were identified and are discussed below.

6.2.1 Content

Content of an institutional repository is the most important factor that shows the success of a repository. It is the core of what constitutes a repository. From the literature authors such as Markey and others (2007), Thibodeau (2007), and McGill (2011), argue that content is the backbone of any repository. Therefore, repository managers must strive to populate the repositories through mandatory policies as voluntary deposits are low and produce little content. It has also been found in the literature that content recruitment which brings in a lot of content in

the repository could only come about because of self-deposits but academics are not happy about self-archiving.

At the UCT Institutional Repository, the contents have grown significantly from 310 photographs since the first digitization project of 2001, to the current total of 4063. According to the Head of the Digitization Unit the numbers of the contents have improved tremendously because of the use of the software DigiTool. At the moment the repository contents include digital collections with 1893 materials, 1199 theses and dissertation and 971 finding aids (last updated on the 30th September 2011).

6.2.2 Use

The second factor that has been identified as a success factor is use and it can be divided into three sub-categories: *number of users*, *type of content used* and *nature of use* (Harnad and McGovern, 2009). The university student body, mainly the postgraduate students and the academic staff, use repositories. Other users identified are researchers outside the university community and stakeholders who are interested in particular fields. Whenever users search the Web, they may retrieve articles from a university or research institutional repository.

The most common type of content used by the above mentioned repository users are peer-reviewed journal articles (preprints and post prints by faculty researchers), bibliographic references (metadata only), books, sections and chapters, conference and workshop papers, theses and dissertations, unpublished reports and working papers, content-packaged learning objects, multimedia and audio-visual materials, patents, and datasets. Datasets may be in the form of spreadsheets, photographs, representations of artwork, diagrams, and charts. The main indicator of use is through downloads.

In the UCT Institutional Repository, the factor of use is not yet available.

6.2.3 Submissions

The rate and number of submissions, which leads to growth of the repository, is the third factor to consider in judging if the repository is a success. It has been discussed that institutional repository managers should produce submission policies which include the following aspects:

depositors; the type of materials that will be deposited; the type of materials that can be deposited and the format; policies stating how the contents will be packaged; the level of moderation for checking deposits, e.g. checking intellectual copyright for legal materials deposited; and quality check of sound quality of audiovisual materials.

At UCT, the presently available Library policy deals with submission of electronic theses and dissertations by students and digital collections. A new policy should be formulated as soon as possible as for the repository to run effectively it should be guided by a policy.

6.2.4 Advocacy

Advocacy deals with building an informed awareness among stakeholders for which a core message and ethos are essential. The institutional repository must be embedded in the institution. Promotion can be achieved by using different approaches, making sure all the stakeholders hear the message and existence of the institutional repository from as many people and through as many channels as possible. This must be an ongoing process, which must be repeated as many times as possible for everyone, new or old, to hear. Reasons for setting up an institutional repository must be articulated as clearly as possible. The key to success in bringing cultural change will be achieved when making researchers, who are the main depositors of content, understand why the institutional repository has been established and what their role is in the repository. At UCT, this has not yet been done.

6.2.5 Support

Support is another factor to consider in judging the success of an institutional repository. Without support from the university community and the stakeholders, the institutional repository will never be successful. In the UCT Institutional Repository, the factor of support has been achieved through the funding from the Carnegie Foundation of New York. The extent to which it will be supported by the university community remains to be seen.

6.2.6 Influence

Influence refers to providing assistance to other institutions that may want help when they start establishing their institutional repository.In the UCT Institutional Repository, this is already

happening, as the head of Digital Initiatives Unit conducts workshops and consults people who need help with digitization.

6.2.7 Interoperability

This refers to where the exchange of digital objects between repositories is possible and it is usually the responsibility of organizations to ensure that the workflow for data exchange is carefully designed. In simpler terms, systems that comply with the interoperability standards are able to communicate with each other in such a way that data found in one system is comprehensible and usable by another. Metadata plays a role in the interoperability of systems or the institutional repository (Nabe, 2010:30). This means that for digital repositories to achieve interoperability and to exchange digital objects between them, they first need to provide data and/or metadata, which they can understand to their partners. To achieve interoperability, "the parties need to agree on the structure, the syntax and the semantics of the data and/or metadata objects, and they need to either comply with common formats or to provide a means for metadata mapping and/or for object conversion." (Aschenbrenner & Kaiser, 2005: 25). The common format for repositories is the Open Archives Initiative-Protocol for Metadata Harvesting (OAI-PMH) proposed by OAI. For OAI compliance, a repository must expose and allow dissemination of unqualified Dublin Core Metadata, which is discussed above in chapter 2.10.7 (Aschenbrenner & Kaiser, 2005). The institutional repository at UCT will be interoperable.

Although UCT selected proprietary software for its institutional repository, it meets the needs and requirements of open access, which refers to the free global access to scholarly information published electronically. The UCT institutional repository will be an open access initiative in spite of being based on proprietary software.

6.3 Summary of the research questions and findings that answer the questions

Below is a synopsis of the research questions and the findings of the study. The study objectives were founded on the following questions;

- What are the steps involved in the development, establishment and implementation of an institutional repository and what measures have to be taken to overcome identified challenges?
- How does the UCT institutional repository compare and contrast with that at UP?
- To what extent will the process followed at UCT in developing an institutional repository lead to a resource that meets with criteria for success?

The research project intended to answer the questions above and with regard to the first question. These steps were: identifying role players, addressing the issue of resources, evaluating software, formulating a policy for the Institutional Repositories, restructuring the library to accommodate change and licensing. This has been discussed in chapter 5.2.

The second question was addressed in chapter 5.3 where the UCT and UP institutional repositories were compared. The comparison showed that UP is far ahead of UCT. Its two repositories were established in 2000 and 2006 respectively. The contents in these two repositories is more than the one in the UCT repository. The comparison also showed that the two UP repositories are recognized internationally as they are both registered with international indexing services and repository directories such ROAR, OpenDOAR, Google, and OAISTER. The ETD repository is also registered with NDLTD while UPSpace is registered with the DSpace community. The UCT repository is still at its infancy stage and will only do the registration in future after it has been launched.

The last question, which attempted to establish the extent to which the process followed at UCT in developing an institutional repository, will lead to a resource that meets the criteria for success, was found to be positive. The project team members followed the steps identified in chapter 5.2 to establish their repository and at the time of writing, were at the stage of restructuring the library to accommodate change. It was found that some of the changes were that staff members involved in the establishment of the repository had already been trained and had also visited American universities to benchmark how to set up institutional repositories. Some staff were being redeployed from other UCT Libraries Units to the DIU. The launch of the institutional repository to the UCT community will mark the end of the process of establishing a

repository hence will have been successful in following the process. They will then register it so that UCT research can be visible to the whole world.

6.4 Recommendations

The recommendations below arise from the observations and the analysis from the study. They are based on the critical areas and challenges that were encountered. It is in this context therefore that the following interventions are proposed below.

6.4.1 Ensuring that the unit moves to a bigger location

The area where the DIU is located has limited space. It would be advantageous if the UCT Library Management could move this unit to a bigger location.

The latest developments with regard to space at the DIU are that some of the tables that are used by the UCT Libraries IT department have been taken out. This allowed the DIU to rearrange their furniture and create a wider space. However, this is not enough and a larger location is required.

6.4.2 Employ more members of staff

In order to improve productivity and reduce the backlog of materials that need to be digitised, the UCT Libraries management should redeploy some of the librarians in the UCT library who have expertise on digitisation to the Digital Unit. If they do not have enough staff members in the library, they should employ new members of staff with appropriate qualifications to assist the existing members who are unable to cope with the backlog.

6.4.3 Buy more equipment

The unit could improve productivity and the backlog of material to be digitised by purchasing additional hardware. This includes flatbed scanners, film and negative scanners, more desktop computers, and high-resolution cameras.

6.4.4 In-service training

Another recommendation is that some members of staff, especially the junior staff, could be trained on the job, thus receiving in-service training. They could organise workshops with other librarians in the region, or cluster like CALICO and share experiences and ideas concerning how they do things in their libraries. This would be helpful in reducing the amount of money the library would spend on sending the junior staff on outside training. A further possibility could be to invite skilled technicians to run seminars or workshops as sending people for training or for short courses may be expensive and may exacerbate the shortage of staff.

6.4.5 Market the repository

It is recommended that the DIU should market the repository as soon as the launch date is announced. Marketing should continue after the launch in order to make the UCT community aware of the existence of the repository and the services they offer. Marketing should be done in order to raise the information awareness level of users and stakeholders. They could use some of the marketing strategies proposed in chapter two (cf. 2.9.4). This includes knowing the audience, content of the marketing message, and delivering the message. They must also ensure that the message is appealing and addresses the needs of the community they are targeting.

6.5 Overall conclusion

It has been argued that an institutional repository could provide a valuable service in terms of the availability of information. This study set out to provide an overview of how UCT established its repository. It has been shown in the literature that implementation of institutional repositories is a value-added service that can be provided to an institution such as UCT. These services include increased access, more visibility, and preservation of digital information. When compared with UP, UCT established its institutional repository in a different manner. Other institutions started by using a consultancy, while UCT used its own experts in the field of institutional repository. This has been discussed in detail in chapter 5 (cf.5.2.1). Proprietary software was used as opposed to open source software. The disadvantage is that DigiTool is not flexible and if there is

a problem, Ex Libris will provide its own expertise. However, UP uses open source software and has trained its own staff.

The literature provided valuable guidelines and information regarding the development of the repository. UCT used some of these guidelines such as formulating policy, identifying role players and choosing software to use (cf.5.2). It is essential that changes in technology and new trends be monitored on a continual basis. Technology obsolescence is generally regarded as the greatest technical threat to ensuring continued access to digital material. Institutional repositories are playing an important role in digital preservation and long-term curation. It therefore enables effective usage of digital formats that require detailed planning and budgeting. It is also essential that detailed risk assessments be carried out to determine the effect of changes in media, upgrades, new software and hardware.

Although universities worldwide have embraced the establishment and implementations of institutional repositories, there are still challenges that they need to address. These challenges include convincing the academic researchers to deposit their research material in the institutional repository (cf. 2.9.1) and the challenge of marketing the institutional repository to the university community so that everyone in the institution is aware of the institutional repository existence. Advocacy has been proposed as a means to assist institutional repositories in educating their communities about their existence on campus and teach them about their benefits. A number of success indicators have been discussed in detail in order to highlight what an institutional repository should do to be successful.

6.6 Suggestions for further research

As the research progressed, a number of ideas for further research emerged. These are listed below:

- Investigating strategies for how UCT's DIU would market the Institutional Repository to the community it will serve. In the literature, repositories are faced with the challenge of academics who are not willing to deposit their research materials in the institutional repositories. As a result, institutional repositories remain empty after being established.

Content recruitment is the biggest challenge faced by institutional repository managers and can be addressed by marketing.

- An exploratory study which targets the UCT academic staff only, where one explores and investigates the perceptions, attitudes and responses of academic authors and researchers with respect to open access (OA) publishing and institutional repositories, especially towards the UCT Institutional Repository. Obtaining UCT academics' views about the new scholarly publishing model brought about by the open access movement through institutional repositories would be beneficial.

- The third suggestion would be to investigate the use of institutional repositories in South Africa focusing on who uses them. This is important for those institutions that have established an institutional repository in the last five to ten years. This particular question could be narrowed to the particular provincial area in the form of a comparative analysis. For example, comparing institutional repositories in the Western Cape, Eastern Cape or Gauteng or Free State.

References

Alvi, W.A. 2004. Electronic scholarly publishing and libraries. *2nd International CALIBER*. 11-13 February 2004. INFLIBNET Centre, Ahmedabad, New Delhi. Available: http://shodhganga.inflibnet.ac.in/dxml/bitstream/handle/1944/341/04cali_33.pdf?sequence=1 [2011, August 3].

Amaral, M. 2008. *Institutional Repositories, Open Source Options, and Libraries*. Available: www.infosherpas.com/ojs/index.php/openandlibraries/article/view/11/20. [2011, March 23].

Anbu, K.P.J. 2006. *Institutional repositories: time for African universities to consolidate the digital divide*. Available: http://www.ascleiden.nl/Pdf/elecpublconfanbu.pdf [2010, November 15].

Arms, W. Y. 2000. *Digital Libraries*. Cambridge, MA: The MIT Press.

ARWU. 2011. *Academic ranking of world universities*. Available: http://www.shanghairanking.com/ARWU2011.html. [2011, June 19]

Asamoah –Hassan, H. 2007. Embracing electronic scholarly publishing in Africa: The Kwame Nkrumah University of Science and Technology (KNUST) Library. Kumasi, Ghana as a case study. First PKP Conference in Vancouver, Canada 11-13 July 2007. Available: http://scholarlypublishing.blogspot.com/2007/07/embracing-electronic-scholarly.html [2011, March 5].

Asamoah-Hassan, H. 2010. Alternative scholarly communication: management issues in a Ghanaian university. *Library Management*. 31(6): 420-426. Available: http://www.emeraldinsight.com/journals.htm?articleid=1864903. [2011, March 5].

Aschenbrenner, A & Kaiser, M. 2005. *White paper on digital repositories*. reUSE digital master files of printed material eContent Project No. 11173. Available: http://www.uibk.ac.at/reuse/docs/reuse-d11_whitepaper_10.pdf. [2011, August 3].

ASSAF website. Available: http://www.assaf.org.za. [2011, March 5].

Ayris, P., Maccoll, J. & Pinfield, S. 2002. *SHERPA: Securing a Hybrid Environment for Research Preservation and Access*. Revised proposal in response to JISC Circular 1/02: Focus on Access to Institutional Resources (FAIR) Programme. Available: http://www.sherpa.ac.uk/documents/proposal.pdf. [2011, March 7].

Babbie, E. & Mouton, J. 2001. *The practice of social science research*. South African edition. Cape Town: OUP.

Bailey, C.W. 2006. *SPEC kit 292: institutional repositories July 2006*. Washington, DC: Association of Research Libraries.

Ball, A. 2010. *Preservation and curation in institutional repositories*. Edinburgh, UK: Digital Curation Centre. Available: http://www.era.lib.ed.ac.uk/bitstream/1842/3381/1/Ball%20irpc-report-v1.3.pdf [2011, August 30].

Bellinger, et al. 2002. The state of digital preservation: an international perspective. *Proceedings of the institute for Information Science Council on Library and Information Resources.* 24-25 April. Washington, D.C. Avaialble: http://www.clir.org/pubs/reports/pub107/pub107.pdf. [2012, March 24]

Berlin Declaration. 2003. *Open access to knowledge in the sciences and humanities.* Available:http://www.zim.mpg.de/openaccess-berlin/berlindeclaration.html. [2010, November 15].

Bethesda statement on open access publishing. 2003. Available: http://www.earlham.edu/~peters/fos/bethesda.htm.[2010, November 15].

BOAI. 2002. *Budapest open access initiative*. Available: http://www.soros.org/openaccess/read.shtml. [2010, November 15].

BOAI, 2011 http://www.soros.org/openaccess/list_signatures). [2010, November 15]

Bogdan, R. 1992. *Qualitative research for education: an introduction to theory and methods*. Boston : Allyn and Bacon.

Bogdan, R. & Taylor, S.J. 1975. *Introduction to qualitative research methods.* New York: John Wiley.

Borgman, C. L. 1999. What are digital libraries? Competing versions. *Information Processing and Management.* 33: 227 – 243.

Bothma, T., Pienaar, H. and Hammes, M. 2008. Towards open scholarship at the University of Pretoria. *16th BOBCATSSS Symposium 2008: Providing Access to Information for Everyone*. Available: http://edoc.hu-berlin.de/conferences/bobcatsss2008/bothma-theo-271/PDF/bothma.pdf. [2011, July 7].

Brophy, P. & Frey, F. 2006. *Digital curation centre externally moderated reflective self-evaluation: Report*. Available: http://ierepository.jisc.ac.uk/198/1/dcc_evaluation_report_final.pdf. [2011, August 22].

Campbell-Meier, J. 2008. Creating narratives for project management and assessment: case studies on institutional repository development. PhD Thesis. University of Hawaii. Available: http://scholarspace.manoa.hawaii.edu/handle/10125/4177. [2011, March 12].

Caplan, P. 2003. *Metadata fundamentals for all librarians*. Chicago: American Library Association.

Carnegie Mellon University. 2011. *Scholarly communication at Carnegie Mellon University.* Available: http://www.cmu.edu/scholarly-communications/. [2011, September 1].

Carpenter, N. 2008. Tune it up: creating and maintaining the institutional repository revolution. *Open and Libraries Class Journal.* 1(1). Available: http://www.infosherpas.com/ojs/index.php/openandlibraries/article/view/13/23 [2011, May 8].

Carr, L. & Brody, T. 2007. Size isn't everything: sustainable repositories as evidenced by sustainable deposit profiles. *Lib Magazine* 13(7/8). Available: http://www.dlib.org/dlib/july07/carr/07carr.html. [2011, March 3].

Chan, L. 2004. Supporting and enhancing scholarship in the digital age: the role of open-access institutional repositories. *Canadian Journal of Communication.* 29(3). Available: http://eprints.rclis.org/bitstream/10760/5648/1/Chan_CJC_IR.pdf. [2010, November 10].

Chan, L. & Swan, A. 2003. *Open access repositories: open access scholarly information source book - practical steps for implementing open access.* Available: http://www.openoasis.org/index.php?option=com_content&view=article&id=137&Itemid=333. [2011, March 3].

Christian, G.E. 2008. Issues and challenges to the development of open access institutional repositories in academic and research institutions in Nigeria. *Proceedings ELPUB 2008 Conference on Electronic Publishing - Toronto, Canada.* June 2008. Available: http://idl-bnc.idrc.ca/dspace/bitstream/10625/36986/1/127792.pdf. [2010, November 15].

Conway, P.1990. *Archival preservation in a nationwide context.* American Archivist. 53(2):204-222.

Corletey, A. 2011. Institutional repositories for open access: the Ghanaian experience. *Proceedings of the 14th International Symposium on Electronic Theses and Dissertations.*13-17 September 2011, Cape Town, South Africa. Available: http://dl.cs.uct.ac.za/conferences/etd2011/papers/etd2011_corletey.pdf. [2011, October 31]

Cox, Glenda. 2010. Open access. Handout for workshop a during the open access week 2010. 22-26 October. Unpublished.

Crow, R. 2002. *The case for institutional repositories: a SPARC position paper, scholarly publishing and academic resources coalition.* Washington: DC. Available: http://www.arl.org/sparc/ir/ir_final_release_102.pdf. [2010, November 15].

Cullen, R. and Chawner, B. 2008a. Institutional repositories in tertiary institutions: access, delivery, and performance. In J. R. Griffiths & J. Craven (Eds.), *Access, delivery, and performance in tertiary libraries without walls: A festschrift to celebrate the work of Professor Peter Brophy.* 113-144. London: Facet.

Cullen, R. and Chawner, B. 2009. Institutional repositories and the role of academic libraries in scholarly communication. *Proceedings of A-LIEP (Asia-Pacific Association of Library and Information Education and Practice*. March 2009. Tuskuba. Available: http://researcharchive.vuw.ac.nz/bitstream/handle/10063/866/conferece%20paper.pdf?sequence=2. [2011, February 12].

Cullen, R. & Chawner, B. 2010. Institutional repositories: assessing their value to the academic community. *Performance Measurement and Metrics.* 11(2): 131–147. Available: http://www.emeraldinsight.com/journals.htm?articleid=1871182&show=html. [2011, August 30].

De Beer, J.A. 2005. Open access scholarly communication in South Africa: current status, significance, and the role for national information policy in the national system of innovation, South Africa. MPhil Thesis. Department of Information Science, Stellenbosch University. Available: http://eprints.rclis.org/archive/00003922/01/DeBeerJennifer.MThesis.Final.pdf [2010, November 15]

Deegan, M. and Tanner, S. 2002. *Digital futures: strategies for the information age*. London: Library Association Publishing.

Department of Public Service & Administration. 2006. Policy on free and open source software use for South African government. Available: http://www.info.gov.za/view/DownloadFileAction?id=94490. [2012, March 27]

DigiTool. 2011. Available: http://www.digitool.uct.ac.za. [2011, January 25].

Dhliwayo, J. 2004. Developing a fibre optic backbone for Africa. A paper presented at ITU Telecom Africa NEPAD Council. Cairo. http://www.corning.com/docs/corporate/discovery_center/innovation_library/2004/ntr100763.pdf. [2012, March 20]

DINI Working Group. 2010. *DINI certificate document and publication services.* Electronic Publishing. Available: http://edoc.hu-berlin.de/series/dini-schriften/2010-3-en/PDF/dini-zertifikat-2010-3-en.pdf. [2011, August 2].

Drake, M.A. 2004. Institutional repositories: hidden treasures. *Searcher*.12(5) May. Available: http://www.infotoday.com/searcher/may04/drake.shtml. [2011, March 16].

DRIVER. 2008a. *DRIVER Guidelines 2.0: Guidelines for content providers- exposing textual resources with OAI – PHM.* Available: http://www.driversupport.eu/documents/DRIVER_Guidelines_v2_Final_2008-11-13.pdf [2011, August 3].

DRIVER. 2008b. *Digital repository infrastructure vision for European research.* Information for institutional staff. Available: http://www.driver-support.eu/personnel.html.[2011, May 5].

Dunlop, J. & Hart, L. 2005. *Digitization projects at the University of Cape Town Libraries.* Available: www.innovation.ukzn.ac.za/InnovationPdfs/No30pp32-42Dunlop&Hart.pdf. [2011, May 5].

Ex Libris Group. 2011. Available: http://www.exlibrisgroup.com/category/DigiToolOverview. [2011, March 16].

Foster, N.F. & Gibbons, S. 2005. Understanding faculty to improve content recruitment for institutional repositories. *D-Lib Magazine.* 11(1). Available: http://www.dlib.org/dlib/january05/foster/01foster.html. [2010, December 6].

Garett, J. & Waters, D. 1996. *Report of the task force on archiving of digital information.* Available: www.clir.org/pubs/reports/pub63watersgarrett.pdf. [2011, October 3]

Gibbons, S. 2004. *Establishing an institutional repository.* Chicago: American Library Association.

Gray, E. 2009. *University of Pretoria adopts open access mandate*. Available: http://blogs.uct.ac.za/blog/gray-area/2009/05/25/a-mandate-at-the-university-of-pretoria. [2011, May 3].

Gray, E. 2007. A critique of research dissemination policy in South Africa, with recommendations for reform. CPS International Policy Fellowship Programme, Central European University Centre for Policy Studies and the Open Society Institut.Budapest. Available: http://blogs.uct.ac.za/gallery/215/eve.gray.pdf. [2012, March 30].

Greenstein, D. 2000. *Digital library federation draft strategy and business plan.* Digital Library Federation. Available: http://www.diglib.org/about/stragic.htm#archs . [2010, November 17].

Guterman, L. 2008. Harvard faculty adopts open-access requirement. *The Chronicle of Higher Education News Blog.* Available: http://chronicle.com/news/article/3943/harvard-faculty-adopts-open-access-requirement. [2011, May 3].

Hammes, M. 2010. *UPeTD celebrates 10 years of success.* Available: http://www.library.up.ac.za/openup/docs/UPeTD10.pdf. [2011, May 22].

Harnad, S. 2001. The self archiving initiative: freeing the refereed research literature online. *American Scientist Open Access Forum.* Available: http://cogprints.org/1642/1/nature4.htm [2011, March 3].

Harnad, S. & McGovern, N. 2009. Institutional repository success is dependent upon mandates. *Bulletin of the American Society for Information Science and Technology.* 35(4): 1-5. Available: http://onlinelibrary.wiley.com/doi/10.1002/bult.2009.1720350410/pdf. [2011, March 27].

Hitchcock, S., & others. 2007. Digital preservation service provider models for institutional repositories. *D-Lib Magazine.* 13(5/6). Available: www.dlib.org/dlib/may07/hitchcock/05hitchcock.html. [2010, November 13].

Hitchcock, S. & others. 2005. *Preservation for institutional repositories: practical and invisible.* Available: http://www.ukoln.ac.uk/events/pv-2005/pv-2005-final-papers/033.pdf. [2010, November 13].

Hockx-Yu, H. 2006. Digital preservation in the context of institutional repositories. *Program Electronic Library and Information Systems.* 40(3): 232-243. Available: http://www.emeraldinsight.com/journals.htm?articleid=1563488. [2010, November 15].

Hodge, G. & Carroll, B.C. 1999. Digital electronic archiving: the state of the art and the state of the practice. A report sponsored by the International Council for Scientific and Technical

Information Policy committee and CENDI. Available: http://www.icsti.org/IMG/pdf/digarch99_Main-2.pdf. [2011, April, 11].

Jain, P., Bentley, G and Oladiran, M.T. 2009. The role of the institutional repository in digital scholarly communications. *Proceedings of the African Digital Scholarship and Curation Conference.* 12-14 May 2009, CSIR Conference Centre: Pretoria. Available: www.ais.up.ac.za/digi/docs/jain_paper.pdf. [2011, September 4].

Jantzi, L. 2009. *Challenges and solutions: faculty content and institutional repositories.* Available: http://leannajantzi.ca/documents/Jantzi594FINAL.pdf. [2010, October 11].

Jensen, M. 2006. Open access: lowering the costs of international bandwidth in Africa. APC Issue Papers. October. Available: http://www.apc.org/fr/system/files/open_access_EN.pdf. [2011, October 19].

JISC Repository Net. 2011. *Open archives initiative-protocol for metadata harvesting (OAI-PMH).* Available: www.rsp.ac.uk/documents/briefing-papers/technical-%20OAI_PMH.pdf [2011, August 02].

Johnson, G. J. 2008. In the kingdom of the blind: successfully implementing institutional repositories in the United Kingdom and the Sherpa partnership experience. *New Review of Academic Librarianship.* 13(1-2):13-33. Available: http://www.informaworld.com/smpp/content~db=all~content=a792024743pdf. [2011, May 2].

Johnson, G. J. 2007a. *SHERPA Road show delegate briefing document.* Available: http://www.sherpa.ac.uk/documents/sherpaplusdocs/sherpa-roadshow-handout-final.pdf [2011 March 15].

Johnson, G. J. 2007b. *Voices of authority: advocating institutional repositories for successful cultural change.* SHERPA Roadshow, University of St Andrews. http://www.sheraa.ac.uk/documents/StAndrews-talk1-GJJ-apr2007.pdf [2011, March 16].

Johnson, G. J. 2007c. *Into the fire: the SHERPA partnership's 7½ pillars of advocacy.* SHERPA Road show. De Montfort University. Available: http://www.sherpa.ac.uk/documents/dmu-v2gjj-talk3-mar.pdf [2011, March 16].

Johnson, K. & Magusin, E. 2005. *Exploring the digital library: a guide for teaching and learning.* San Francisco: Jossey Bass-Wiley.

JISC. 2005a. *Digital repositories: helping universities and colleges.* Available: www.jisc.ac.uk/uploaded_documents/JISC-BP-Repository(HE)-v1-final.pdf. [2010, November 15].

JISC. 2005b. *Open access briefing paper.* Available: www.jisc.ac.uk/uploaded_documents/JISC-BP-OpenAccess-v1-final.pdf. [2011, March 16].

Jones, M. & Beagrie, N. 2002. *Preservation management of digital materials: a handbook.* York: Digital Preservation Coalition. Available: http://www.dpconline.org/graphics/intro/definitions.html. [2011, March 5].

Jones, R., Andrew, T. & MacColl, J. 2006. *The institutional repository.* Oxford, U.K: Chandos. Available: http://www.rsp.ac.uk. [2011, March 17].

Kerlinger, F. N.1964. *Foundations of behavioural research.* Available: http://www.librarything.com/author/kerlingerfredn [2010, April 13].

Kim, J. 2007. Motivating and impeding factors affecting faculty contribution to institutional repositories. *Journal of Digital Information.* 8(2). Available: http://journals.tdl.org/jodi/article/viewArticle/193/177. [2010, November 3].

Kim, J. 2010. Faculty self-archiving: motivations and barriers. *Journal of the American Society for Information Science and Technology.* 61(9): 1909 -1922. Available: http://onlinelibrary.wiley.com/doi/10.1002/asi.21336/pdf. [2011, August 19].

Kingsley, D. 2008. Those who don't look don't find: disciplinary considerations in repository advocacy. *OCLC Systems & Services: International digital library perspectives.* 24(4): 204-218.

Kling, R. 2004. The internet and unrefereed scholarly publishing. In *Annual Review of Information Science and Technology.* 38(1): 591-631. Edited by B Cronin. Medford, NJ: Information Today Inc. Available: http://onlinelibrary.wiley.com/doi/10.1002/aris.1440380113/pdf. [2011, July 5].

Kling, R. & McKim, G. 1999. Scholarly communication and the continuum of electronic publishing. *Journal of the American Society for Information Science.* 50(10):890–906. Available: http://onlinelibrary.wiley.com/doi/10.1002/%28SICI%291097-4571%281999%2950:10%3C890::AID-ASI6%3E3.0.CO;2-8/pdf. [2011, April 11].

Lamptey, R. B. & Corletey. 2011. *Enhancing institutional repositories (IR) in Ghana* in open access to STM information: trends, models, and strategies for libraries. Anthi Katsirikou. Ed. De Gruyter Saur: Berlin, Boston. Available: http://dspace.knust.edu.gh:8080/jspui/bitstream/123456789/716/1/Lamptey.pdf [2011, March 3].

Lynch, C.A. 2003. Institutional repositories: essential infrastructure for scholarship in the digital age. *Libraries and the Academy.* 3(2):327-336. Available: DOI: 10.1353/pla.2003.0039. [2010, November 15].

Macha, A. & De Jager, K. 2011. Two South African institutional repositories: a comparative overview. *Proceedings of the 14th International Symposium on Electronic Theses and Dissertations.* 13-17 September 2011, Cape Town, South Africa. Available. http://dl.cs.uct.ac.za/conferences/etd2011/papers/etd2011_macha.pdf. [2011, October 31]

Mark, T. & Shearer, K. 2006. Institutional repositories: a review of content recruitment strategies. *World Library and Information Congress: 72nd IFLA General Conference and Council.* 20-24 August 2006. Seoul, Korea: IFLA. Available: http://www.ifla.org/IV/ifla72/index.htm [2011, March, 22].

Markey, K. and others. 2007. *Census of institutional repositories in the United States: MIRACLE Project Research Findings.* Washington, DC: Council on Library and Information Resources, 2007. Available: http://www.clir.org/pubs/abstract/pub140abst.html [2011, March 16].

Markland, M. & Brophy, P. 2005. SHERPA Project evaluation: final report. Manchester: Centre for Research in Library & Information Management. Available: http://www.sherpa.ac.uk/documents/SHERPA_Evaluation.pdf_ [2011, March 17].

McGill, L. 2011. *Digital repositories: infokit- JISC.infoNet.* Available: http://www.jiscinfonet.ac.uk/infokits/repositories/digital-repositories-infokit.pdf [2011, September 20].

McLennan, J. & Mark, T. 2006. SPARC and CARL commend new policy for public access to outputs of Canadian research. *Canadian Institutes of Health Research Introduces Draft Policy Implementing Broad Access to Publicly-Funded Research with Minimal Restrictions.* Available : http://www.arl.org/sparc/advocacy/06-1128.shtml. [2011, May 03]

Merton, R.K &. Kendall, P.L. 1946. The focused interview. *American Journal of Sociology.* 51 (6): 541-557. May 1946. The University of Chicago Available: http://www.jstor.org/stable/2770681. [2012, March 19]

Moahi, K. 2009. Institutional repositories: towards harnessing knowledge for African development. *1st International Conference on African Digital Libraries and Archives (ICADLA-I).* 1 - 3 July 2009. Addis Ababa, Ethiopia. Available: http://wiredspace.wits.ac.za/bitstream/handle/10539/8950/35%20Moahi.pdf?sequence=1 [2011, March 10].

Monash University. 2011. *eResearch centre.* Available: http://www.monash.edu.au/eresearch/about/whatis.html. [2011, October 10].

Mouton, J. & Marais, H.C.1990. *Basic concepts in the methodology of the social sciences.* 2nd ed. Pretoria: HSRC.

Nabe, J.A. 2010. *Starting, strengthening, and managing institutional repositories: a how-to-do-it-manual.* New York: Neal-Schuman.

Nachmias, C. & Nachimias D. 1996. *Research methods in the social science.* 5th ed. London: Arnold.

National Information Standards Organization. 2004. *Understanding metadata*. Bethesda: NISO Press. Available: http://www.dspace.up.ac.za/metadata.htm. [2012, March 24]

Neuman, L. Ed. 1994. *Social research methods: qualitative and quantitative approaches.* 2nd edition. Boston: Allyn & Bacon.

Nfila, R.B. 2007. *Academic libraries' support for e-learning: initiatives and opportunities: the case of University of Botswana library*. Available: http://www.ais.up.ac.za/digi/docs/nfila_paper.pdf. [2011, July 11].

Noor, K.B.M. 2008. Case study: a strategic research methodology. *American Journal of Applied Sciences* 5(11):1602-1604. Available: http://findarticles.com/p/articles/mi_7109/is_11_5/ai_n28551353/pg_2/. [2011, April 23].

Olivier, E. 2010. O*pen access and the evolving scholarly communication environment workshop.* 1-2 September 2010.Available: https://repository.up.ac.za/upspace/bitstream/2263/14788/1/Olivier_Electronic%282010%29.pdf [2011, May 14].

Olivier, E. & Louw, I. 2011. The dot on the i: reflections of a postgraduate student and repository manager on the etd process at the University of Pretoria. *Proceedings of the 14th International Symposium on Electronic Theses and Dissertations*.13-17 September 2011, Cape Town, South Africa. Available: http://dl.cs.uct.ac.za/conferences/etd2011/papers/etd2011_olivier.pdf. [2011, October 31]

Olivier, E. 2011. Advocacy for the open access mandates implementation @ the University of Pretoria: a case study. *Proceedings of the 13th Annual LIASA Conference*. 3-7 October 2011. East London. South Africa. Available: http://conference2011.liasa.org.za/presentations/olivier_ppt_in_pdf.pdf [2012, March 26].

Oppenheim, C. 2008. Electronic scholarly publishing and open access. *Journal of Information Science,* 34(4):577-590. Available: DOI:10.1177/0165551508092268. [2011, April 28].

OpenDOAR, 2009. http://www.opendoar.org/find.php. [2011, October 20].

Organ, M & Deveson, L. 2010.*Two models for institutional repository success*. Available: http://.www.*digitalcommons.bepress.com/cgi/viewcontent.cgi?filename.valatalk.doc*. [2011. February 15].

Organ, M. K. 2006. Download statistics - what do they tell us? the example of research online. The Open Access Institutional Repository at the University of Wollongong. Australia 2006. *D-Lib Magazine* 12 (11). Available: http://ro.uow.edu.au/asdpapers/44. [2011, March 14].

Pappalardo, K. & Fitzgerald, A. 2007. A *guide to developing open access through your digital repository*. Available: http://www.oaklaw.qut.edu.au/node/32. [2011, February 15].

Park. J. & Qin. J. 2007. Exploring the willingness of scholars to accept open access: a grounded theory approach. *Journal of Scholarly Publishing*. 38(2):55-84. Available: http://muse.jhu.edu/journals/scp/summary/v038/38.2park.html [2011, April 11].

Pasqui, V. 2007. *Digital preservation and open access archives: persistent access to open access digital assets*. Available: http://www.digitalpreservationeurope.eu/publications/briefs/open_archives_pasqui.pdf [2011, July 3].

Pelizzari, E. 2005. Harvesting for disseminating: open archives and the role of academic libraries. *The Acquisitions Librarian*. 33/34: 35-51. Available: http://pdfserve.informaworld.com/170778_751318378_907946278.pdf. [2011, June 12].

Peters, D. 2003. Giving form to the future: challenge to preservation in the digital age. *South African Archives Journal*. 43: 79-88. Available: EbscoHost Academic search premier: http://www.lib.uct.ac.za/datahosts.htm. [2011, September 4].

Phillips, M. E. 1999. The national library of Australia: ensuring long-term access to online publications. *Journal of Electronic Publishing*. 4(4).Available: DOI:10.3998/3336451.0004.405 [2011, June 2].

Pickton, M. & McKnight, C. 2007. Is there a role for research students in an institutional repository? Some repository managers' views. *Journal of Librarianship and Information Science.* 39(3):153-161.Available: http://lis.sagepub.com/content/39/3/153 [2011, July 7].

Pienaar, H. & Van Deventer, M. 2007. Capturing knowledge in institutional repositories - playing leapfrog with giraffes. *Proceedings of the73rd IFLA General Conference and Council Knowledge Management pre-conference workshop at the World Library and Information Congress (WLIC).* 17 August 2007. Available: http://researchspace.csir.co.za/dspace/bitstream/10204/1107/1/Pienaar_2007.pdf [2010, November 15].

Pinfield, S. 2004. Self-archiving publications. in: Gorman, G E and Rowland, F. Eds. *International yearbook of Library and Information Management 2004-2005: Scholarly publishing in an electronic era.* London: Facet.118-145. Available: http://eprints.nottingham.ac.uk/archive/0000 . [2011, June 29].

Pinfield, S. 2005. A mandate to self-archive? the role of open access institutional repositories. *Serials*, 18(1): 30-34. [2011, March 4].

Prosser. C.D. 2003. Institutional repositories and open access: the future of scholarly communication. *Information Services and Use*. 23(2-3): 167-170. Available: http://iospress.metapress.com/content/b0t9y69afjhylw9y/fulltext.pdf.[2010, December 3].

Prosser, C.D. 2005 a. Open access: the future of scholarly communication. *Association of Portuguese Libraries, Archives and Documentations.* 1: 6-20. Available: http://redalyc.uaemex.mx/pdf/385/38500302.pdf. [2011, March 20]

Prosser. C.D. 2005 b. *Fulfilling the promise of scholarly communication a comparison between old and new access models.* Available: http://eprints.rclis.org/bitstream/10760/6353/1/mittler_Paper.pdf. [2011, February 15]

QS. 2011. *Quacquarelli symonds world university rankings.* Available: http://www.topuniversities.com/university-rankings/world-university rankings/2011?page=3. [2011, August 20].

Rapp, J. 2009. UCT Libraries strategic plan 2010. Unpublished.

Reitz, Joan. M. 2006. *ODLIS - online dictionary for library and information science*. Available: http://lu.com/odlis/odlis_e.cfm. [2010, November 26].

Rieger, O.Y. 2008. Opening up institutional repositories: social construction of innovation in scholarly communication. *Journal of Electronic Publishing* .11(3). Available: http://dx.doi.org/10.3998/3336451.0011.301. [2010, December 3].

ROMEO. 2011. *Project ROMEO*. Available: http://www.lboro.ac.uk/departments/ls/disresearch/romeo/index.html. [2011, August 3].

SARUA. 2011. Available: http://www.sarua.org/. Southern African Regional Universities Association. [2011, June 14].

Shearer, K. 2006. The CARL institutional repositories project: a collaborative approach to addressing the challenges of IRs in Canada. *Library Hi Tech*, 24(2):165-172. Available: http://www.emeraldinsight.com/journals.htm?articleid=1558871. [2011, March 8].

Shearer. K. 2003. Institutional repositories: towards the identification of critical success factors. *Canadian Journal of Information and Library Science*. 27(3): 250-263. Available. http://cais-acsi.ca/proceedings/2003/Shearer_2003.pdf. [2011, March 8]

Smith, A. 2004. The digital preservation conundrum part 1. *Serial Librarian*. 46 (1/2): 107-113. Available:http://dx.doi.org/10.1300/J123v46n01_12. [2010, November 23].

Smith, B. 2003. Preserving tomorrow's memory: preserving digital content for future generations. A newsletter of the IFLA: Preservation and Conservation. *International Preservation News*. (29):4-9. May. Available: http://archive.ifla.org/VI/4/news/ipnn29.pdf. [2011, October 6].

Smith, I. 2005. An institutional repository for the University of Pretoria. *DITCHE Conference and workshop.* 20 May 2005. Port Elizabeth Nelson Mandela Metropolitan University. Available: www.dspace.up.ac.za/DITCHE.pps. [2012, March 24].

Smith, I. 2008. University of Pretoria digital institutional research repository: business plan. University of Pretoria. Available: http://www.dspace.up.ac.za. [2012, March 23]

Smith, I & Pienaar, H. 2007. Adapt or die: energizing library staff and academics through the development of digital repositories. *Proceedings of the73rd IFLA General Conference and Council Knowledge Management pre-conference workshop at the World Library and Information Congress (WLIC).* 17 August 2007. Available: https://www.up.ac.za/dspace/handle/2263/3325 [2011, May 2].

Smith, J.G. 2006. A longitudinal study of the information communication process among a defined group of scientists in South Africa. PhD thesis: UCT. Unpublished.

Song, S. 2011. *Africa undersea cables*. Available: http://manypossibilities.net/african-undersea-cables/. [2012, March 30].

SPARC Europe. 2007. SHERPA receives SPARC europe award for outstanding achievements in scholarly communications. Available: http://www.sparceurope.org/press_release/SHERPA.html_. [2011, March 17].

SPARC. 2011. Avaialble: http://www.sparceurope.org/. [2012, March 28]

Stake, R. 1995. *The art of case research*. Thousand Oaks, CA: Sage.

Stevenson, V. & Hodges, S. 2008. Setting up a university digital repository: experience with DigiTool. 24(1):48-50. *Emerald Group Publishing Limited.* Available: http://www.emeraldinsight.com/journals.htm?articleid=1674217&show. [2011, May 6].

Swan, A. 2006. The culture of open access: researchers' views and responses. in Jacobs, N. Ed. Open access: key strategic technical and economic aspects. Chapter 7. Chandos, Oxford: 52-59. http://eprints.ecs.soton.ac.uk/12428/01/asj7.pdf. [2011, March 29]

Swan, A. and Awre, C. 2006. Linking UK repositories: technical and organisational models to support user-oriented services across institutional and other digital repositories: scoping study report. Available:

http://www.jisc.ac.uk/uploaded_documents/Linking_UK_repositories_report.pdf [2011, March 17].

Swan, A. & Brown, S. 2004. Authors and open access publishing. *Learned Publishing.* 17 (3): 219-226. Available: http://www.ingentaconnect.com/content/alpsp/lp/2004/00000017/00000003/art0000710 [2010, November 15].

Swan, A. & Brown, S. 2005. *Open access self–archiving: an author study.* Available: http://www.keyperspectives.co.uk/openaccessarchive/reports/Open%20Access%20II%20(author%20survey%20on%20self%20archiving)%202005.pdf [2010, November 15].

Tellis, W. 1997. Introduction to case study. *The Qualitative Report Serial* 3(2). Available: http://www.nova.edu/ssss/QR/QR3-2/tellis1.html. [2011, March 7].

Thibodeau, K. 2007. If you build it, will it fly? criteria for success in a digital repository. *Journal of Digital Information.* 8(2). Available: http://journals.tdl.org/jodi/article/view/197/174. [2011, April 24].

Thomas, G. 2007. Evaluating the impact of the institutional repository, or positioning innovation between a rock and a hard place. *New Review of Information Networking.* 13(2):133-146. Available: http://dx.doi.org/10.1080/13614570802105992. [2011, March 24].

Treloar, A. 2004. *Building an institutional research repository from the ground up: the ARROW architecture experience.* Available: http://ausweb.scu.edu.au/aw04/papers/refereed/treloar/paper.html. [2011, March 18].

Troll Covey, T. 2011. Recruiting content for the institutional repository: the barriers exceed the benefits. *Journal of Digital Information.* 12(3): 1-18. Available: http://journals.tdl.org/jodi/article/view/2068. [2011, August 17].

Umit, H. 2005. *Case study.* Available: http://christophe.benavent.free.fr/IMG/pdf/CaseStudy_UCL_2005.pdf. [2011, May 14].

University of Cape Town Libraries. 2011. Available: http://www.lib.uct.ac.za. [2011, January 18].

University of Cape Town. 2011. Available: http://www.uct.ac.za. [2011, January 18].

University of Pretoria. 2011. Available: http://upetd.up.ac.za/UPeTD.htm. [2011, June 02].

University of South Carolina. 2009. *Digital traditions dictionary of folklore McKissick museum.* Institute of Museum and Library Services. Available: http://www.digitaltraditions.net/html/Glossary.cfm#D. [November, 2010 30].

Van der Merwe, A. 2008. Development and implementation of an institutional repository within a science, engineering and technology environment. MPhil Dissertation. Department of Informatics, University of Pretoria, South Africa. Available: http://upetd.up.ac.za/thesis/available/etd08282008090829/unrestricted/dissertation.pdf [2011, May 6].

Van Deventer, M. & Pienaar, H. 2008. *South African repositories: bridging knowledge divides.* Available: http://www.ariadne.ac.uk/issue55/vandeventer-pienaar/. [2011, May 6].

Van de Sompel, H. & Lagoze, C. 2000. The Santa Fe convention of the open archives initiative. *D-Lib Magazine.* 6(2). Available:http://dx.doi.org/10.1045/february2000-vandesompel-oai. [2011, June 18].

Walters, T.O. 2007. Reinventing the library: how repositories are causing librarians to rethink their professional roles. *Portal: Libraries and the Academy.* 7(2):213-25. Available: http://smartech.gatech.edu/handle/1853/14421. [2011, October 15]

Ware, M. 2004. Institutional repositories and scholarly publishing. *Learned Publishing.* 17:115-124. Available: http://www.ingentaconnect.com/content/alpsp/lp/2004/00000017/00000002/art00006. [2010, December 28].

Westell, M. 2006. Institutional repositories in Canada institutional repositories: proposed indicators of success. *Library Hi Tech.* 24(2): 211-226. Available: http://www.emeraldinsight.com/journals.htm?issn=0737-8831&volume=24&issue. [2011, February 6].

Wheatley, P. 2004. Institutional repositories in the context of digital preservation. Available: http://www.dpconline.org/docs/DPCTWf4word.pdf. [2011, March 17].

Whitehead, D. 2005. Repositories: what is the target? an arrow perspective. *New Review of Information Networking*. 11(1):123-134. Available: http://dx.doi.org/10.1080/13614570500293553. [2010, October 17].

Yin, R. 1994. *Applications of case study research*. Beverly Hills, CA: Sage Publishing.

Yin, R. 2003. *Case study research: design and methods*.3rd ed. Beverly Hills, CA: Sage.

Yin, R. 2009. *Case study research: design and methods.* 4th ed. Los Angeles: Sage.

Young, J.R. 2002. Superarchives could hold all scholarly output: online collections by institutions may challenge the role of journal publishers. *Chronicle of Higher Education*, 48(43). Available: http://chronicle.com/free/v48/i43/43a02901.htm. [2011, June 18].

APPENDICES: APPENDIX A

University of Cape Town
Rondeberg Flat 302
7 Main Road
7700
RONDEBOSCH
Cape Town, South Africa

31st May 2011

The Director
University of Cape Town Libraries
Private Bag X3, 7701 RONDEBOSCH
South Africa

Dear Sir/Madam

RE: PERMISSION FOR CONDUCTING RESEARCH IN THE LIBRARY

I am an MPhil candidate in the Department of Information and Library Studies, University of Cape Town, my supervisor is A/Prof Karin de Jager. I would like to request your permission to conduct research in the library in the **Digital Initiatives Unit** as part of my studies. My research will be exploring the *Establishment and Implementation of a digital repository at an academic institution in South Africa: challenges faced and how they were overcome: case of UCT.* It will also identify challenges that are encountered and based on the investigation; recommendations will be made to improve the situation.

The study will be conducted by means of counting documents, archival records, participant observation, questionnaires and interview with some of the library staff who are directly involved with the development of the UCT institutional repository and knowledgeable about the issues taking place. The UCT's Faculties of Education and Research Ethics Board have already approved the study; it will be carried out from 2nd of June to the 30th June 2011.

I hope my application will be considered.

Yours faithfully

Annah. S. Macha

APPENDIX B

The administrative documents that were found while being attached at the Digital Initiatives Unit (DIU) which were important to the research project are:

- UCT Libraries ETD Policy
- Scanning Policy for ETD
- UCT Libraries Strategic Plan for the years 2010 - 2014
- Submitting your thesis: guide for students
- Metadata fields for still images and theses
- Proposal Form
- Copyright page
- CD cover sheet
- Delivery Form for materials to the Digital Initiatives Unit (DIU)
- Minutes from the Digital Content Committee
- Minutes from the UCT Libraries Management Team

APPENDIX C

Questions for the Head of the DIU

Interview Guide Questions arranged according to objectives

1. **Which factors influenced University of Cape Town Library in establishing an institutional repository?**
    - Which individuals, organizations or resources played key roles in the founding of the IR?
    - What obstacles were encountered in the founding of the IR?
    - How were these obstacles addressed?
2. **What resources and structures are being put in place in developing the institutional repository?**
    - Did you carry out a resource audit to find out what resources you already had in preparation for the IR establishment?
    - Which resources were required to develop the IR?
    - Which of these resources are necessary for the ongoing operation of the IR?
3. **To what extent is the institutional repository an open access initiative?**
    - Does the software you chose [Digitool] support open access?
    - Can contents in Digitool be harvested by OAIster?
    - Will content from DSpace and EPrint repositories be simply transferable to Digitool? [e.g. contents from the Computer Science and Law Space repositories]
    - Is metadata in Digitool different from metadata in DSpace concerning the objects? If so explain.

4. **Are the departmental digital repositories going to be brought together under the UCT institutional repository?**

   - Do you have a strategy on how you are going to bring the departmental digital repositories together?
   - Why do you want to bring them together?
   - When will they be brought together?

5. **What are the criteria that could be used to evaluate the success of an IR?**

   - What would you say are the factors that can be used to judge if UCT IR is successful?
   - How is the IR assessed? Are you using any tools to assess the success of the IR?
   - What are they?
   - How would you define IR success?
   - Do you have any examples of the benefits the IR has produced?

6. **What challenges have been encountered in the digital initiatives unit in the establishment of the institutional repository?**

   - What challenges have you faced in implementing the IR?
   - What are the future challenges and opportunities you see for the IR
   - What lessons or challenges are faced by IR developers that may impact development?

7. **How have the challenges been dealt with?**

   - What is going well?
   - What is not going well?
   - How do you manage the challenges?
   - What would you do differently?
   - Why would you do these things differently?

Questions for the manager of the Manuscripts and Archives Department in the UCT Libraries

1. **Which factors influenced University of Cape Town Library in establishing an institutional repository?**

   - What were the events that led to the founding of the IR?
   - What was the biggest motivating factor?
   - What obstacles were encountered in the founding of the IR?
   - What are the sources of funding for the IR?
   - Ask Lesley Hart about the artefacts they have at the Department of Manuscripts and Archives. Ask her what artefacts do they have and how are they stored.
   - To what extent will images will be digitised & made available in the IR
   - What challenges have you faced in implementing the IR?
   - What is going well?
   - What is not going well?
   - How do you manage the challenges?
   - What would you do differently?
   - Why would you do these things differently?

Questions for the Head of Information Technology in the UCT libraries

2. **What resources and structures are being put in place in developing the institutional repository?**

   - What are the sources of funding for the IR?

- Which resources were required to develop the IR?
- Which of these resources are necessary for the ongoing operation of the IR?
- What challenges have you faced in implementing the IR?

7. **How have the challenges been dealt with?**
   - What is going well?
   - What is not going well?
   - How do you manage the challenges?
   - What would you do differently?
   - Why would you do these things differently?

Questions for the technician in the DIU

- What is your role in the IR?
- Does UCT IR have a regular tracking system of hits to see how frequently the IR is accessed?
- How can the hits downloaded from UCT IR be tracked?
- UCT chose proprietary software over DSpace. What is, in your opinion, the main benefit of DigiTool over DSpace?
- Have you encountered any obstacles in your use of DigiTool?
- What challenges have you faced in implementing the IR?
- What is going well in the IR?
- What is not going well?
- How do you manage the challenges?

Questions for Senior Curator in the Manuscripts and Archives department of the UCT Libraries

- You joined the IR and found it having been in its process of development towards establishment. What is it that you found missing?
- What have you has done towards the development of the IR?
- What still remains to be done?
- How have you contributed to the IR?
- What is, in your opinion, the goal of UCT's IR?
- What is going well in the IR?
- What challenges have you faced in implementing the IR?
- What is going well in the IR?
- What is not going well?
- How do you manage the challenges?

APPENDIX D

UCT LIBRARIES

MANUSCRIPTS AND ARCHIVES DEPARTMENT

SCALE OF CHARGES FOR REPROGRAPHIC AND OTHER SERVICES 2011

Payment is required in advance.

Photocopies

	UCT STAFF & STUDENTS	NON-UCT USERS	POSTAL REQUESTS
1 – 50 copies	A4 R0.50 per copy A3 R1.00 per copy	A4 R1.00 per copy A3 R2.00 per copy	A4 R2.00 per copy A3 R4.00 per copy
50 + copies	A4 R1.00 per copy A3 R1.60 per copy	A4 R2.00 per copy A3 R4.00 per copy	A4 R4.00 per copy A3 R8.00 per copy
Colour photocopies	A4 R10.00	A4 R15.00	

There is a minimum charge of R100.00 for postal requests.

Print-outs from microfilm

	UCT STAFF & STUDENTS	NON-UCT USERS
A4	R1.00	R2.00
A3	R2.00	R4.00

Digital images (Research purposes)

	UCT STAFF & STUDENTS	NON-UCT USERS
High res scans	R25.00	R50.00
Quick flat scans	R15.00	R30.00

Reproduction fees for images in our collections

This is not a copyright fee. Where copyright exists with UCT, it will be charged separately. Where it is held elsewhere, the onus is on the user to obtain the consent of the copyright holder prior to publication. In all cases, the pictures may be used once only and UCT Libraries must be acknowledged as the source. This charge is additional to the cost of supplying a copy of the image.

NEWSPAPERS & MAGAZINES	SOUTH AFRICA	INTERNATIONAL
Standard	R250.00	R500.00
NGOs	R150.00	R300.00
Magazine covers	R800.00	R1,600.00

BOOKS & SCHOLARLY JOURNALS	SOUTH AFRICA	INTERNATIONAL
Standard (incl school textbooks)	R400.00	R800.00
Educational (excl school textbooks)	R150.00	R300.00
Covers	R800.00	R1,600.00

TELEVISION DOCUMENTARIES/ FEATURE FILMS	SOUTH AFRICA	INTERNATIONAL
	R1,000.00 per image, up to 10 seconds	R1,500.00 per image, up to 10 seconds
NGOs	R500.00 per image, up to 10 seconds	R1,000.00 per image, up to 10 seconds

ADVERTISEMENTS & COMMERCIAL USE	SOUTH AFRICA	INTERNATIONAL
Books & magazines	R1,000.00	R1,500.00

Reproduction fees for literary manuscripts

	SOUTH AFRICA	INTERNATIONAL
	R300.00 per page	R600.00 per page

Copying of photographs and collections material on Library premises

- Prior permission must be requested for the use of digital or other cameras.
- Reproduction fees will be charged in respect of each image copied for publication.
- There will be a charge per hour or part thereof for commercial photographers. This is negotiable.

Certified copies of documents

R60.00 per document

Methods of payment:

- Cash
- Cheque: Only guaranteed cheques accepted, made payable to *University of Cape Town*
- Direct deposit into University of Cape Town bank account: Bank details as follows:
  Standard Bank, Riverside Mall, Rondebosch, 7700, Cape Town
  Account number: 07 150 3854
  Swift Address SBZAZAJJ
  Branch: 50 09
  Please quote "Manuscripts & Archives" to enable the receipting cashier to track our department when the money arrives.

Credit card: If payment by credit card is required, please ask us to send you the necessary form